Short Questions

Unit 1:

1. What is the definition of Data Science?

2. How does Datafication contribute to Data Science?

3. What is Statistical Inference in Data Science?

4. What are the key components of Statistical Modelling in Data Science?

5. How does Overfitting impact Statistical Modelling in Data Science?

6. What is the significance of understanding Populations and Samples in Data Science?

7. What are the fundamental Data Types in R programming?

8. How is R-Environment Setup performed for Data Science projects?

9. What are the primary steps involved in Programming with R for Data Science projects?

10. What role do Basic Data Types play in R programming for Data Science?

11. Why is understanding the Current Landscape of Perspectives important in Data Science?

12. How does Statistical Inference contribute to decision-making in Data Science projects?

13. What is the significance of Data Science in addressing Big Data challenges?

15. What distinguishes Data Science from traditional statistical analysis16. How does Data Science address the challenge of Datafication?

17. What are the core objectives of Data Science projects?

18. How does Statistical Inference contribute to the credibility of insights derived from data analysis?

19. What role does Probability Distributions play in Statistical Modelling?

21. What are the primary data types used in R programming?

22. How does Data Science contribute to evidence-based decision-making?

23. What distinguishes Data Science from traditional data analysis approaches?

24. How does understanding populations and samples contribute to the validity of statistical analysis in Data Science?

25. What are some common pitfalls to avoid in statistical modelling in Data Science projects?

26. How does Data Science address the challenge of data quality in real-world applications?

27. What are the key considerations when selecting statistical models for Data Science projects?

28. How does Data Science help businesses gain a competitive advantage?

29. What role does exploratory data analysis (EDA) play in Data Science projects?

30. How does Data Science contribute to predictive analytics in healthcare?

31. What are some ethical considerations in Data Science projects?

33. What are the challenges associated with handling unstructured data in Data Science projects?

34. How does Data Science contribute to personalized recommendation systems in e-commerce?

35. What role does Data Science play in climate modelling and environmental research?

36. How does Data Science contribute to supply chain optimization in logistics and transportation?

37. What are some emerging trends in Data Science and its applications38. How does Data Science contribute to predictive maintenance in manufacturing industries?

39. What are the implications of Data Science for personalized healthcare and precision medicine?

40. How does Data Science contribute to customer segmentation and targeted marketing strategies?

41. How does Data Science contribute to predictive analytics in energy consumption and resource management?

42. What are the key challenges in implementing Data Science solutions in real-world applications?

43. How does Data Science contribute to urban planning and smart city initiatives?

44. What role does Data Science play in sentiment analysis and social media monitoring?

45. How does Data Science contribute to predictive analytics in cybersecurity?

46. How does Data Science contribute to natural disaster prediction and mitigation efforts?

47. How does Data Science contribute to fraud detection in healthcare insurance claims?

48. What are the implications of Data Science for personalized learning and education?

49. How does Data Science contribute to wildlife conservation and biodiversity monitoring?

50. What are the future prospects of Data Science and its impact on society and technology?

Unit 2:

51. What are the different types of data attributes?

52. How do you describe attributes based on the number of values they can assume?

53. What distinguishes nominal attributes from ordinal attributes?

54. How do you measure the central tendency of data?

55. What is the range of a dataset?

56. How do quartiles describe the dispersion of data?

57. How is variance calculated, and what does it signify?

58. What is the significance of standard deviation in data analysis?

59. How does the interquartile range (IQR) describe data variability?

60. What graphical displays are commonly used for visualizing data distributions?

61. How do asymmetric attributes differ from symmetric attributes?

62. What distinguishes discrete attributes from continuous attributes?

63. How does Data Science utilize nominal attributes in data analysis?

64. What statistical measures are commonly used for describing data variability?

65. How does Data Science handle ordinal attributes in data analysis?

66. How are binary attributes represented and analysed in Data Science?

67. How do Data Science practitioners handle numeric attributes in data analysis?

68. What graphical displays are suitable for visualizing numeric attributes in data analysis?

69. What is the role of nominal attributes in descriptive statistics?

70. How do Data Science practitioners identify outliers in data analysis?

71. How do asymmetric attributes differ from symmetric attributes?

72. What distinguishes discrete attributes from continuous attributes?

73. How does Data Science utilize nominal attributes in data analysis?

74. What statistical measures are commonly used for describing data variability?

75. How does Data Science handle ordinal attributes in data analysis?

76. How are binary attributes represented and analysed in Data Science?

77. How do Data Science practitioners handle numeric attributes in data analysis?

78. What graphical displays are suitable for visualizing numeric attributes in data analysis?

79. What is the role of nominal attributes in descriptive statistics?

80. How do Data Science practitioners identify outliers in data analysis?

81. How do asymmetric attributes differ from symmetric attributes?

82. What distinguishes discrete attributes from continuous attributes?

83. How does Data Science utilize nominal attributes in data analysis?

84. What statistical measures are commonly used for describing data variability?

85. How does Data Science handle ordinal attributes in data analysis?

86. How are binary attributes represented and analysed in Data Science?

87. How do Data Science practitioners handle numeric attributes in data analysis?

88. What graphical displays are suitable for visualizing numeric attributes in data analysis?

89. What is the role of nominal attributes in descriptive statistics?

90. How do Data Science practitioners identify outliers in data analysis?

91. How do asymmetric attributes differ from symmetric attributes?

92. What distinguishes discrete attributes from continuous attributes?

93. How does Data Science utilize nominal attributes in data analysis?

94. What statistical measures are commonly used for describing data variability?

95. How does Data Science handle ordinal attributes in data analysis?

96. How are binary attributes represented and analysed in Data Science?

97. How do Data Science practitioners handle numeric attributes in data analysis?

98. What graphical displays are suitable for visualizing numeric attributes in data analysis?

99. What is the role of nominal attributes in descriptive statistics?

100. How do Data Science practitioners identify outliers in data analysis?

Unit 3:

101. What is the purpose of creating and naming vectors in data science?

102. How does vector arithmetic contribute to data manipulation tasks?

103. What is vector sub setting, and how is it useful in data analysis?

104. How are matrices created and named in data science workflows?

105. What is matrix sub setting, and how does it contribute to data analysis?

106. How do arrays differ from matrices in data science applications?

107. What is the role of factors in data analysis, and how are they summarized?

108. How do ordered factors differ from nominal factors, and why is this distinction important?

109. What techniques are used for comparing ordered factors in data analysis?

110. How are data frames introduced in data analysis, and what is their role?

111. How are data frames subsetted, and why is this operation important in data analysis?

112. How are data frames extended, and what benefits does this provide in data analysis?

113. How are data frames sorted in data analysis workflows, and why is sorting important?

114. What is the role of lists in data analysis workflows, and how are they created?

115. How are named lists created and accessed in data analysis workflows?

116. What operations can be performed to manipulate list elements in data analysis?

117. How are lists merged or concatenated in data analysis workflows?

118. What techniques are available for converting lists to vectors in data analysis?

Additionally, specific list elements can be extracted and combined into a vector using indexing or sub setting operations. These techniques allow analysts to transform list-based data structures into vector formats suitable for further analysis, visualization, or modelling tasks in statistical programming environments like R.

119. How do lists differ from vectors and matrices in data analysis workflows?

120. How are lists converted to data frames, and why is this transformation valuable in data analysis?

121. What is the significance of sub setting data frames in data analysis, and how is it performed?

122. How are factors summarized in data analysis, and why is this process important?

123. What distinguishes ordered factors from nominal factors, and why is this distinction relevant?

124. How are ordered factors compared in data analysis, and what insights can be gained from this comparison?

125. What is the role of data frames in data analysis workflows, and how are they extended?