



QP CODE: 24018588



24018588

Reg No :

Name :

M Sc DEGREE (CSS) EXAMINATION, APRIL 2024
Fourth Semester
M Sc COMPUTER SCIENCE (DATA ANALYTICS)
Elective - CA860402 - BUSINESS DATA ANALYTICS

2020 ADMISSION ONWARDS

3E87A378

Time: 3 Hours

Weightage: 30

Part A (Short Answer Questions)

*Answer any **eight** questions.*

Weight 1 each.

1. Describe the importance of business analytics.
2. Write short note on framework for business analytics.
3. Differentiate between Discrete and Continuous data.
4. Find the Median for the given data: 90,94,53,68,79,94,53,65,87,90,70,69,65,89,85,53,47,61,27,80
5. Calculate the Correlation Coefficient for the following data: 82,23,59,94,70,26,32,83,87,94,32
6. Define events and its probabilities with an example.
7. Define hypothesis testing.
8. Define Hadoop.
9. Describe union and intersection by using MapReduce.
10. Differentiate between SQL and HiveQL.

(8×1=8 weightage)

Part B (Short Essay/Problems)

*Answer any **six** questions.*

Weight 2 each.

11. Define business analytics. Explain the need for business analytics.
12. Explain the various modelling techniques and algorithms used in business data analysis.





13. Calculate the Quartile Deviation and its coefficient of the following data: 24,32,,46,48,39,42,28,25,26,24,38
14. Write short note on a). Heatmap b). Dashboards.
15. a) Define the probability density function of a random variable. b)Let X is a random variable which denotes the number of heads obtained when a coin is tossed 3 times. Find the probability distribution of the random variable X and mean and variance of the distribution.
16. The probability that a bulb produced by the factory will fuse after 150 days of use is 0.05. Find the probability that out of five such bulbs, a). None of them will fuse after 150 days of use. (Hint: $x=0$) b). Not more than one will fuse after 150 days of use. (Hint $x \leq 1$, $x=0$ or $x=1$) c). More than one will fuse after 150 days. d). Atleast one will fuse after 150 days.
17. Differentiate between Spark and tensor-flow systems.
18. Explain the use of Apache Pig and Pig Latin.

(6×2=12 weightage)

Part C (Essay Type Questions)

Answer any **two** questions.

Weight 5 each.

19. Explain the applications of business analytics.
20. Calculate the mean, variance and standard deviation for the following distribution: Class Frequency(f_i)
30-40 3 40-50 7 50-60 12 60-70 15 70-80 8 80-90 3 90-100 2
21. From the following data of marital status and if customer defaults on mortgage: a). What is the probability that a customer defaults on his or her mortgage b). What is the probability that a customer does not default on his or her mortgage c). Calculate the probability that a randomly selected customer is married and that the customer does not default on his or her mortgage. d). What is the probability that a randomly selected customer is single and that the customer does not default on his or her mortgage. e). Compute the probability that a customer defaults on his or her mortgage given that the customer is married.
22. Explain in detail MapReduce algorithm and explain its working with an example.

(2×5=10 weightage)

