

E

Reg. No. :

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Question Paper Code: 52R01

M.E. DEGREE EXAMINATION, APRIL 2019

Second Semester

Computer Science and Engineering

15PCS201 - DATA SCIENCE AND BIG DATA ANALYTICS

(Regulation 2015)

Duration: Three hours

Maximum: 100 Marks

Answer ALL Questions

PART - A (5 x 20 = 100 Marks)

1. (a) Classify EDW and big data in terms of analytics . CO1- U (20)
Or
(b) What are the Three Characteristics of Big data, and the main considerations in processing Big data? CO1- U (20)
2. (a) Explain the core deliverables for the stake holders. CO2- U (20)
Or
(b) Describe the Analytic Plan for Churn Prediction for Retail Banking. CO2- U (20)
3. (a) Use lm () and ggplot () package to do the ANOVA. CO3-App (20)
Or
(b) Create a function to calculate the pooled variance, which is used in the student's statistic. CO3-App (20)
4. (a) Illustrate Logistic Regression with example. CO4- U (20)
Or
(b) Describe the Association Rules and Apriori Algorithm with an Example. CO4- U (20)

5. (a) Explain In-database Analytics. List the differences between SQL and relational databases? CO5- U (20)

Or

- (b) Highlight the features of Hadoop and explain the functionalities of Hadoop cluster? CO5- U (20)
-