

M.C.A. III SEMESTER – BATCH-B2 LAB-TEST-01

DWDM Laboratory – DATA PREPROCESSING

MAX.MARKS: 10 DATE: 13-SEP-2019

DURATION: 1.5 Hrs.

Input Data File: Train.csv

Convert this file into a .csv file and import it into R environment.

- 1) Display the 5 number summary of Age of passengers
- 2) Visualize the 5 number summary of Age using box plot
- 3) Find the ticket and name, fare of the passenger with the lowest fare.
- 4) Find the total number of family members of passenger with name “Andersson, Mr. Anders Johan”
- 5) Tabulate mean fare for each passenger class.
- 6) Create a contingency table, tabulating sex & passenger class. Infer if more number of males or females were in 1st class.
- 7) Use the chi square test to find the association between Pclass and survived. Infer the result
- 8) Tabulate the total number of children (Age < 18) who survived.
- 9) Draw side-by-side boxplots of the age of passengers and survival.