Program: BBA (Computer Application) Semester: III   **SET: B**
Program (Specific): BBACA07   **Course Type: CC**
Class: S.Y.B.B. A(CA)
Max.Marks: 70
Name of the Course: Big Data Analytics
Course Code: 23-BBACA235   **Time: 2:30 Hrs**
Paper: -

Instructions to the candidate:

1) *There are 4 sections in the question paper. Write each section on separate page.*
2) *All Sections are compulsory.*
3) *Figures to the right indicate full marks.*
4) *Draw a well labelled diagram wherever necessary.*

SECTION: A

Q1) Answer the following .   **[1*10= 10 marks]**

a) _____answer the question "What has happened?"

   i)   Descriptive analytics   ii) Predictive analytics   iii) Prescriptive

   analytics   iv) Diagnostic analysis

b) _____ answers the question "What will happen?"

   i)   Descriptive analytics   ii) Predictive analytics   iii) Prescriptive

   analytics   iv) Simple analysis

c) When data are collected in a statistical study for only a subset of all elements of
   interest we are using_____

   i)   Sample   ii) Population   iii) Parameter   iv) Coefficient

d) _____is most important language for Data Science.

   i)   R   ii) Java   iii) C++   iv) C

e) In Statistics, a population consist of_____

   i)   All people living in a Country.   ii) All people living in the city are under

   Study.   iii) All subjects or objects whose characteristics are being studied.

   iv) All people not living in a Country.

f) Collection of objects currently stored in R is called as____

   i)   Workspace   ii) Package   iii) list   iv) taks

g) File Containing R Scripts end with extension_____.

   i) .S   ii) .R   iii) .RP   iv) .SP

h) To select Columns(Vaiables)_____function is used.

   i) Select()   ii) Filter ()   iii) Mutate()   iv) Sumarise()

i) Unsupervised learning makes sense of _____ data without having any predefined

   dataset for its training.

   i) Unlabeled   ii) Labeled   iii) Semi-labeled   iv) Prelabeled

j) Wich is not basic class of object in R.

   i) Character   ii) Number   iii) Integer   iv) String

SECTION: B

Q2) Attempt the following (any 10)   **[2*10= 20 marks]**

   a) What is sampling.
   b) Define Population.
   c) Enlist components of Data Science.
   d) What is Digital Data?
   e) Enlist tools used for Big Data.

f) Define descriptive Analytics.

g) Define head() and tail()

h) What is Big Data?

i) What is supervised machine learning?

j) What is Regression Analysis?

k) What is WEKA?

l) What is Data Visualization?

SECTION:C

Q3) Attempt the following (any 4)   **[4*5= 20 marks]**

a) Explain Data Frame with Syntax and Example.

b) Differentiate between Structed Data and Semi structured Data.

c) Explain the process of data analysis.

d) Explain 5 V's of Big Data in detail.

e) Write installation process of RStudio in detail.

f) Explain types of regression model.

SECTION:D

Q4) Attempt the following (any 4)   **[4*5= 20 marks]**

a) What is histogram? Explain with example of R Program.

b) Write an R program to find the maximum and the minimum value of a given vector.

c) Write an R program to find Sum, Mean and Product of a Vector.

d) Write an R program to create a Data frame which contains details of 5 employees and display the details in ascending order.

e) Explain Loops in R with example.

f) Explain advantages of Machine Learning.

———————————————