Avinashilingam Institute for Home Science and Higher Education for Women

(Deemed to be University under Category 'A' by MHRD, Estd. u/s 3 of UGC Act 1956)
Re-accredited with 'A+' Grade by NAAC. Recognised by UGC Under Section 12B
Coimbatore - 641 043, Tamil Nadu, India

Master's Degree Examination – June 2021
IV Semester

Class : II MBA **Time: 3 Hours**
Major : Master of Business Administration/MBA-IT **Max. Marks: 100**

18MBAC30S/18MBMC31S Data Sciences with R Programming

PART A **10 x 1 = 10**
Choose the Correct Answer

1. Choose the analysis which combines visualization and datasense
   a. Table b. Exploratory Data
   c. Cross Analysis d. All of the above

2. A commonly cited limitation of R is that objects must generally be stored in _____
   a. Physical Memory b. Virtual memory
   c. File System d. External Drives

3. Which algorithm determines thedefinition of the right answer by finding clusters of data?
   a. k-NN b. k-means c. Greedy d. Logical Regression

4. A _____ is an artificial construction where all extraneous detail has been removed or abstracted.
   a. Model b. Data c. Pixel d. All of the above

5. Social network analysis (SNA) is the process of investigating social structures through the use of networks and _____.
   a. Media b. Graph Theory c. Internet d. All of the above

6. _____ analytics is the process of using data analysis and business intelligence to improve efficiency and streamline everyday operations in real time
   a. Operational b. Service c. Risk d. Financial

7. Fitting a model means that you estimate the parameters of the model using the _____ data.
   a. Collected b. Predicted c. Observed d. All of the above

8. _____ refers to the extraction of data from a website through APIs or other tools.
   a. Download b. Scraping c. Hacking d. All of the above

9. A key advantage that R has over many other statistical packages is _____ capability.
   a. Linear b. Exponential c. Graphical d. All of the above

10. _____ isdone toward the beginning of analysis, and data visualization
   a. Data Modelling b. Regression c. Data Collection d. Exploratory Data Analysis

Part B **5 x 6 = 30**

Answer ALL questions

Each answer should not exceed 400 words or two pages

11. a. Briefly discuss the hype behind big data and data science.

(or)

11. b. Discuss briefly on Datafication.

12. a. Briefly explain the new kinds of data that a data scientist has to be comfortable dealing with

(or)

12. b. Explain briefly on Hidden biases of Big Data.

13. a. Briefly explain on some broad generalizations to consider in developing machine learning algorithms.

(or)

13. b. Explain briefly how Naive Bayes classification method works.

14. a. Briefly explain about the Interfaces used in R to connect to the Outside World.

(or)

14. b. Briefly explain how companies extract meaning from the data they have.

15. a. Explain briefly about Case-Attribute Data versus Social Network Data.

(or)

15. b. Explain briefly about Morningside Analytics.

Part C **5 x 12 = 60**

Answer ALL questions

Question No 20. Case is Compulsory

Each answer should not exceed 800 words or four pages

16. a. Explain in detail the Role of the Social Scientist in Data Science.

(or)

16. b. Explain in detail about why in the age of Big Data, classical statistics methods need to be revisited and reimagined in new contexts.

17. a. Explain in detail about the stages of data science process

(or)

17. b. Explain the Philosophy of Exploratory Data Analysis.

18. a. Explain the Three Basic machine learning Algorithms

(or)

18. b. Explain in detailhow to obtain the data required for analysis from the web.

19. a. Explain in detail about the process of getting Data In and Out of R using functions.

(or)

19. b. Explain in detail the Feature extraction and selection steps of machine learning.

20. **Case Study: (Compulsory question)**
Explain in detail about Social Network Analysis from a Statistical Point of View.

*****