

Enrollment No.....



Faculty of Engineering
End Sem (Even) Examination May-2019
CS3ED06 Data Science

Programme: B.Tech.

Branch/Specialisation: CSE

Duration: 3 Hrs.**Maximum Marks: 60**

Note: All questions are compulsory. Internal choices, if any, are indicated. Answers of Q.1 (MCQs) should be written in full instead of only a, b, c or d.

- Q.1 i. The multidisciplinary field that uses scientific method, processes and algorithms to extract knowledge is 1
 (a) Data science (b) Data structure
 (c) Machine learning (d) None of these
- ii. Data science deals with 1
 (a) Mathematics (b) Computer Science
 (c) Both (a) and (b) (d) None of these
- iii. Set of all possible outcomes of a random experiment is called as 1
 (a) Probability (b) Event
 (c) Sample Space (d) None of these
- iv. A family has two children. If one of them is known to be a girl then what is the probability that both are girls 1
 (a) 1/2 (b) 1/3 (c) 1/4 (d) 1
- v. The act of retrieving specific data from unstructured or poorly structured data sources for further processing is called 1
 (a) Data Analysis (b) Implementation
 (c) Data Extraction (d) Transaction
- vi. The purpose of data scientist is to convert 1
 (a) Raw data to information (b) Information to raw data
 (c) Both (a) and (b) (d) None of these
- vii. The representation of information in the form of a chart, diagram, pictures is called as 1
 (a) Data Mining (b) Data extraction
 (c) Data Visualization (d) Data Dictionary

P.T.O.

[2]

| | | | |
|-----|-------|--|----------|
| | viii. | Tableau is a | 1 |
| | | (a) Data mining tool (b) Data Visualization tool | |
| | | (c) Network simulation tool (d) None of these | |
| | ix. | Which among these in python is used to create an object | 1 |
| | | (a) A class (b) A function (c) A method (d) A constructor | |
| | x. | Which operator in python is right associative | 1 |
| | | (a) * (b) = (c) + (d) % | |
| Q.2 | i. | Name the primary components of data science. | 2 |
| | ii. | Differentiate between structured and unstructured data. | 3 |
| | iii. | Explain the users of data science and their hierarchy. | 5 |
| OR | iv. | Explain various data science techniques. | 5 |
| Q.3 | i. | Let we are flipping three coins. If at least one of them is turned up as tail, find the probability that at least two of them is head | 2 |
| | ii. | Suppose that there is a lock and there are 10 keys out of which 1 key is the correct key of the lock and 9 others are not the keys of this lock. We don't know which out of 10 keys is the correct key is. At a time one key is used to open the lock and if it is not the correct key of the lock then it is thrown away and next key is tried. What is the probability that the lock is opened by the sixth key? | 8 |
| OR | iii. | Suppose there are 3 players A, B and C and A has already reached to final. B and C are in semi-final and one who wins in semi-final is going to fight with A in the final. The probability that B wins in semi-final is 1/3 and probability that C wins in semi-final is 2/3. If A fights with B in final, then probability that A wins is 2/3 and if A fights with C in final then probability that A wins is 1/3. If A wins the final, then what is the probability that A fought with B in final. | 8 |
| Q.4 | i. | Differentiate between data analysis and data analytics. | 3 |
| | ii. | Explain with the help of steps how the data science solves the business problem. | 7 |
| OR | iii. | Describe any three basic tools of data science. | 7 |
| Q.5 | i. | Explain the process of data visualization. | 4 |

[3]

| | | | |
|-----|------|---|----------|
| | ii. | Describe the basic principles of data visualization. | 6 |
| OR | iii. | Name and describe at least four tools of data visualization. | 6 |
| Q.6 | | Attempt any two: | |
| | i. | List the characteristics of python which promote it as a data science tool. | 5 |
| | ii. | Explain the following python libraries | 5 |
| | | (a) SciPy (b) PyBrain | |
| | iii. | Describe the scope of data science project management. | 5 |

Marking Scheme CS3ED06 Data Science

| | | | |
|-----|-------|--|----------|
| Q.1 | i. | The multidisciplinary field that uses scientific method, processes and algorithms to extract knowledge is | 1 |
| | | (a) Data science | |
| | ii. | Data science deals with | 1 |
| | | (c) Both (a) and (b) | |
| | iii. | Set of all possible outcomes of a random experiment is called as | 1 |
| | | (c) Sample Space | |
| | iv. | A family has two children. If one of them is known to be a girl then what is the probability that both are girls | 1 |
| | | (b) 1/3 | |
| | v. | The act of retrieving specific data from unstructured or poorly structured data sources for further processing is called | 1 |
| | | (c) Data Extraction | |
| | vi. | The purpose of data scientist is to convert | 1 |
| | | (a) Raw data to information | |
| | vii. | The representation of information in the form of a chart, diagram, pictures is called as | 1 |
| | | (c) Data Visualization | |
| | viii. | Tableau is a | 1 |
| | | (b) Data Visualization tool | |
| | ix. | Which among these in python is used to create an object | 1 |
| | | (d) A constructor | |
| | x. | Which operator in python is right associative | 1 |
| | | (b) = | |
| Q.2 | i. | Primary components of data science. Four components 0.5 mark for each | 2 |
| | | (0.5 mark * 4) | |
| | ii. | Differentiate between structured and unstructured data. 1 mark for each difference | 3 |
| | | (1 mark * 3) | |
| | iii. | Five users of data science and their hierarchy. 1 mark for each | 5 |
| | | (1 mark * 5) | |
| OR | iv. | Data science techniques. 1 mark for each | 5 |
| | | (1 mark * 5) | |

| | | | |
|-----|------|---|----------|
| Q.3 | i. | Let we are flipping three coins. If at least one of them is turned up as tail, find the probability that at least two of them is head | 2 |
| | ii. | What is the probability that the lock is opened by the sixth key? | 8 |
| OR | | Stepwise marking | |
| | iii. | If A wins the final, then what is the probability that A fought with B in final. | 8 |
| | | Stepwise marking | |
| Q.4 | i. | Differentiate between data analysis and data analytics. 1 mark for each difference | 3 |
| | | (1 mark * 3) | |
| | ii. | Explain with the help of steps how the data science solves the business problem. Stepwise marking | 7 |
| OR | iii. | Any three basic tools of data science. | 7 |
| | | Marks will be awarded in the basis of description | |
| Q.5 | i. | Process of data visualization. | 4 |
| | ii. | Basic principles of data visualization. 1.5 mark for each principle | 6 |
| | | (1.5 mark * 4) | |
| OR | iii. | Name and describe at least four tools of data visualization. 1.5 marks for each tool | 6 |
| | | (1.5 marks * 4) | |
| Q.6 | | Attempt any two: | |
| | i. | Characteristics of python which promote it as a data science tool 1 mark for each characteristic | 5 |
| | | (1 mark * 5) | |
| | ii. | Explain the following python libraries (a) SciPy (b) PyBrain | 5 |
| | | 2.5 marks 2.5 marks | |
| | iii. | Scope of data science project management. 1 mark for each scope | 5 |
| | | (1 mark * 5) | |
