

# Mock Examination Of Data Science Sep -2020

## Data Science

### Set 2

1	<p>A .....is an application or a set of instructions that scans the Web pages in a systematic and automated manner to categorize the information on the basis of user demand.</p> <ul style="list-style-type: none"><li>a. Web crawler</li><li>b. Web browser</li><li>c. Web engine</li><li>d. Search engine</li></ul>	a
2	<p>– helps to understand the fields in the data that account for the most variance between observations and allow for the processing of a reduced volume of data</p> <ul style="list-style-type: none"><li>a. Feature extraction</li><li>b. Dimensionality reduction</li><li>c. Dimensionality extraction</li><li>d. Feature reduction</li></ul>	b
3	<p>_____shows all individual data points.</p> <ul style="list-style-type: none"><li>a. Box-plot</li><li>b. scatter plot</li><li>c. line plot</li><li>d. pie chart</li></ul>	b
4	<p>Five number summary does not include_____ value.</p> <ul style="list-style-type: none"><li>a. Mean</li><li>b. Median</li><li>c. Mode</li><li>d. Max</li></ul>	c
5	<p>XQuery is a functional query language used to retrieve information stored in ---format.</p> <ul style="list-style-type: none"><li>a. Html</li><li>b. Xml</li><li>c. Uml</li><li>d. Jscript</li></ul>	b

6	<p>In data , information and knowledge triangle, Wisdom is an extrapolative and ....., non-probabilistic process.</p> <ul style="list-style-type: none"> <li>a. non-deterministic</li> <li>b. deterministic</li> <li>c. static</li> <li>d. dynamic</li> </ul>	A
7	<p>Suppose we would like to perform clustering on spatial data such as the geometrical locations of houses. We wish to produce clusters of many different sizes and shapes. Which of the following methods is the most appropriate?</p> <ul style="list-style-type: none"> <li>a) Decision Trees</li> <li>b) Density-based clustering</li> <li>c) Model-based clustering</li> <li>d) K-means clustering</li> </ul>	B
8	<p>Which statements are correct for Akaike Information Criteria(AIC) ?</p> <ul style="list-style-type: none"> <li>1. AIC provides a method for assessing the quality of your model through comparison of related models.</li> <li>2. It's based on the Deviance, but penalizes you for making the model more complicated.</li> <li>3. Much like adjusted R-squared, it's intent is to prevent you from including irrelevant predictors.</li> <li>4. It helps to avoid underfitting.</li> </ul> <ul style="list-style-type: none"> <li>a. Statements 2,3,4 only</li> <li>b. Statements 1,2,3 only</li> </ul>	B

	<p>c. Statements 1,2 only</p> <p>d. Statement 1 only</p>	
9	<p>.....is the process of converting data from one format or structure into another format or structure.</p> <p>a. Data cleaning</p> <p>b. Data curation</p> <p>c. Data transformation</p> <p>d. Data analysis</p>	c
10	<p>The .....comes across only True models.</p> <p>a. Bayesian Information Criteria</p> <p>b. Akaike's Information Criteria</p> <p>c. Minimum Description Length</p> <p>d. Maximum Description Length</p>	a
11	<p>In Xml, which expression Selects nodes in the document from the current node that match the selection no matter where they are</p> <p>a. /</p> <p>b. //</p> <p>c. .</p> <p>d. *</p>	b
12	<p>What is the syntax for to sort the documents in ascending or descending order based on a particular field of document in MongoDB</p> <p>a. db.collection_name.sort({field_key:1 or-1})</p> <p>b. db.collection_name.sort({field_key:1or-1}).find()</p> <p>c. db.collection_name.find().sort({field_key:1or-1})</p> <p>d. db.collection_name.sort({field_key:1or-1})</p>	C
13	<p>The result is a tree-based representation of the observations which is called a .....</p> <p>a. tree plot</p> <p>b. <b>dendrogram.</b></p> <p>c. Density plot</p> <p>d. Scatter plot</p>	B

14	<p>.....is a non-relational DMS, that does not require a fixed schema, avoids joins, and is easy to scale.</p> <ul style="list-style-type: none"> <li>a. SQL</li> <li>b. NoSQL</li> <li>c. Xml</li> <li>d. Html</li> </ul>	b
15	<p>In Cloud computing, PaaS stands for</p> <ul style="list-style-type: none"> <li>a. Public as a Service.</li> <li>b. Private as a Service.</li> <li>c. Platform as a Service.</li> <li>d. Platform as a Section</li> </ul>	c
16	<p>..... Test is used to compare variances of two samples.</p> <ul style="list-style-type: none"> <li>a. Sign test</li> <li>b. T test</li> <li>c. Chi-square test</li> <li>d. F test</li> </ul>	d
17	<p>CouchDB, MongoDB are very popular ..... NoSQL databases for high volume data storage</p> <ul style="list-style-type: none"> <li>a. Key-value Pair Based</li> <li>b. Column-oriented</li> <li>c. Graphs based</li> <li>d. Document-oriented</li> </ul>	d
18	<p>Which one is the false statement given below</p> <ul style="list-style-type: none"> <li>a. HBase is a database built on bottom of the HDFS.</li> <li>b. HBase provides fast lookups for larger tables.</li> <li>c. It provides low latency access to single rows from billions of records.</li> <li>d. HBase internally uses Hash tables and provides random access</li> </ul>	a
19	<p>Select the properties of PCA</p> <ol style="list-style-type: none"> <li>1. A principal component is a normalized linear combination of the original predictors in a data set.</li> <li>2. The PCs are orthogonal.</li> <li>3. The variation present in the PCs decrease as we move from the 1st PC to the last one, hence the importance.</li> </ol>	c

	<p>4. <b>PCA is not sensitive</b> to the relative scaling of the original variables.</p> <ul style="list-style-type: none"> <li>a. Properties 1, 2</li> <li>b. Properties 1, 2, 3</li> <li>c. Properties 1, 2, 4</li> <li>d. Properties 2, 4</li> </ul>	
20	<p>Which is not correct statement for MongoDB?</p> <ul style="list-style-type: none"> <li>a. MongoDB uses binary protocol and custom protocol over TCP/IP.</li> <li>b. In MongoDB, database contains collections and collection contains documents.</li> <li>c. It follows the column-oriented model but data is presented in BSON format.</li> <li>d. MongoDB follows Map/Reduce (JavaScript) creating collection + object-based query language.</li> </ul>	c