| | |
|---|---|
| Which of the following is an example of big data utilized in action today?<br>A. Individual, Unconnected Hospital Databases<br>B. The Internet<br>C. Wi-Fi Networks<br>D. Social Media | D. Social Media |
| What reasoning was given for the following: why is the "data storage to price ratio" relevant to big data?<br>A. Larger storage means easier accessibility to big data for every user because it allows users to download in bulk.<br>B. It isn't, it was just an arbitrary example of big data usage.<br>C. Companies can't afford to own, maintain, and spend the energy to support large data storage unless the cost is sufficiently low.<br>D. Lower prices mean larger storage becomes easier to access for everyone, creating bigger amounts of data for client-facing services to work with. | D. Lower prices mean larger storage becomes easier to access for everyone, creating bigger amounts of data for client-facing services to work with. |
| What is the best description of personalized marketing enabled by big data?<br>A. Marketing to each customer on an individual level and suiting to their needs.<br>B. Being able to obtain and use customer information for groups of consumers and utilize them for marketing needs.<br>C. Being able to use personalized data from every single customer for personalized marketing needs. | C. Being able to use personalized data from every single customer for personalized marketing needs. |
| Of the following, which are some examples of personalized marketing related to big data?<br>A. Facebook revealing posts that cater towards similar interests.<br>B. News outlets gathering information from the internet in order to report them to the public.<br>C. A survey that asks your age and markets to you a specific brand. | A. Facebook revealing posts that cater towards similar interests. |
| What is the workflow for working with big data?<br>A. Extrapolation -> Understanding -> Reproducing<br>B. Theory -> Models -> Precise Advice<br>C. Big Data -> Better Models -> Higher Precision | C. Big Data -> Better Models -> Higher Precision |
| Which is the most compelling reason why mobile advertising is related to big data?<br>A. Mobile advertising benefits from data integration with location which requires big data.<br>B. Since almost everyone owns a cell/mobile phone, the mobile advertising market is large and thus requires big data to contain all the information.<br>C. Mobile advertising allows massive cellular/mobile texting to a wide audience, thus providing large amounts of data.<br>D. Mobile advertising in and of itself is always associated with big data. | A. Mobile advertising benefits from data integration with location which requires big data. |
| What are the three types of diverse data sources?<br>A. Information Networks, Map Data, and People<br>B. Machine Data, Map Data, and Social Media<br>C. Machine Data, Organizational Data, and People<br>D. Sensor Data, Organizational Data, and Social Media | C. Machine Data, Organizational Data, and People |
| What is an example of machine data?<br>A. Weather station sensor output.<br>B. Social Media<br>C. Sorted data from Amazon regarding customer info. | A. Weather station sensor output. |
| What is an example of organizational data?<br>A. Social Media<br>B. Satellite Data<br>C. Disease data from Center for Disease Control. | C. Disease data from Center for Disease Control. |

| | |
|---|---|
| Of the three data sources, which is the hardest to implement and streamline into a model?<br>A. Machine Data<br>B. Organizational Data<br>C. People | C. People |
| Which of the following summarizes the process of using data streams?<br>A. Integration -> Personalization -> Precision<br>B. Big Data -> Better Models -> Higher Precision<br>C. Theory -> Models -> Precise Advice<br>D. Extrapolation -> Understanding -> Reproducing | A. Integration -> Personalization -> Precision |
| Where does the real value of big data often come from?<br>A. Using the three major data sources: Machines, People, and Organizations.<br>B. Combining streams of data and analyzing them for new insights.<br>C. Having data-enabled decisions and actions from the insights of new data.<br>D. Size of the data. | B. Combining streams of data and analyzing them for new insights. |
| What does it mean for a device to be "smart"?<br>A. Collect data and services autonomously.<br>B. Having a specific processing speed in order to keep up with the demands of data processing.<br>C. Must have a way to interact with the user. | A. Collect data and services autonomously. |
| What does the term "in situ" mean in the context of big data?<br>A. Accelerometers.<br>B. Bringing the computation to the location of the data.<br>C. The sensors used in airplanes to measure altitude.<br>D. In the situation | B. Bringing the computation to the location of the data. |
| Which of the following are reasons mentioned for why data generated by people are hard to process? Choose all that apply.<br>A. Skilled people to analyze the data are hard to come by.<br>B. They cannot be modeled and stored.<br>C. Very unstructured data.<br>D. The velocity of the data is very high. | ACD |
| What is the purpose of retrieval and storage; pre-processing; and analysis in order to convert multiple data sources into valuable data?<br>A. Since the multi-layered process is built into the Neo4j database connection.<br>B. To allow scalable analytical solutions to big data.<br>C. Designed to work like the ETL process.<br>D. To enable ETL methods. | B. To allow scalable analytical solutions to big data. |
| Which of the following are benefits of organization-generated data? Choose all that apply.<br>A. Higher Sales<br>B. Better Profit Margins<br>C. High Velocity<br>D. Improved Safety<br>E. Customer Satisfaction | ABDE |
| What are data silos and why are they bad?<br>A. Highly unstructured data. Bad because it does not provide meaningful results for organizations.<br>B. A giant centralized database to house all the data production within an organization. Bad because it hinders opportunity for data generation.<br>C. A giant centralized database to house all the data produces within an organization. Bad because it is hard to maintain as highly structured data. | D. Data produced from an organization that is spread out. Bad because it creates unsynchronized and invisible data. |

| | |
|---|---|
| D. Data produced from an organization that is spread out. Bad because it creates unsynchronized and invisible data. | |
| Which of the following are benefits of data integration? Choose all that apply.<br>A. Monitoring of data.<br>B. Increase data collaboration.<br>C. Adds value to big data.<br>D. Increase data availability.<br>E. Unify your data system.<br>F. Reduce data complexity. | BCDEF |
| Amazon has been collecting review data for a particular product. They have realized that almost 90% of the reviews were mostly a 5/5 rating. However, of the 90%, they realized that 50% of them were customers who did not have proof of purchase or customers who did not post serious reviews about the product. Of the following, which is true about the review data collected in this situation?<br>A. Low Valence<br>B. High Valence<br>C. High Volume<br>D. High Veracity<br>E. Low Volume<br>F. Low Veracity | F. Low Veracity |
| As mentioned in the slides, what are the challenges to data with a high valence?<br>A. Complex Data Exploration Algorithms<br>B. Difficult to Integrate<br>C. Reliability of Data | A. Complex Data Exploration Algorithms |
| Which of the following are the 6 V's in big data?<br>A. Variety<br>B. Valence<br>C. Vision<br>D. Volume<br>E. Velocity<br>F. Value<br>G. Veracity | ABDEFG |
| What is the veracity of big data?<br>A. The size of the data.<br>B. The speed at which data is produced.<br>C. The connectedness of data.<br>D. The abnormality or uncertainties of data. | D. The abnormality or uncertainties of data. |
| What are the challenges of data with high variety?<br>A. The quality of data is low.<br>B. Hard to perform emergent behavior analysis.<br>C. Hard to integrate.<br>D. Hard in utilizing group event detection. | C. Hard to integrate. |
| Which of the following is the best way to describe why it is crucial to process data in real-time?<br>A. More expensive to batch process.<br>B. Batch processing is an older method that is not as accurate as real-time processing.<br>C. Prevents missed opportunities.<br>D. More accurate. | C. Prevents missed opportunities. |
| What are the challenges with big data that has high volume?<br>A. Speed Increase in Processing<br>B. Cost, Scalability, and Performance<br>C. Storage and Accessibility<br>D. Effectiveness and Cost | B. Cost, Scalability, and Performance |

| | |
|---|---|
| Which of the following are parts of the 5 P's of data science and what is the additional P introduced in the slides?<br>A. Platforms<br>B. Purpose<br>C. Process<br>D. Programmability<br>E. Product<br>F. People<br>G. Perception | ABCDEF |
| Which of the following are part of the four main categories to acquire, access, and retrieve data?<br>A. NoSQL Storage<br>B. Traditional Databases<br>C. Web Services<br>D. Remote Data<br>E. Text Files | ABDE |
| What are the steps required for data analysis?<br>A. Regression, Evaluate, Classification<br>B. Classification, Regression, Analysis<br>C. Investigate, Build Model, Evaluate<br>D. Select Technique, Build Model, Evaluate | D. Select Technique, Build Model, Evaluate |
| Of the following, which is a technique mentioned in the videos for building a model?<br>A. Validation<br>B. Evaluation<br>C. Analysis<br>D. Investigation | C. Analysis |
| What is the first step in finding a right problem to tackle in data science?<br>A. Ask the Right Questions<br>B. Define Goals<br>C. Assess the Situation<br>D. Define the Problem | D. Define the Problem |
| What is the first step in determining a big data strategy?<br>A. Collect Data<br>B. Build In-House Expertise<br>C. Organizational Buy-In<br>D. Business Objectives | D. Business Objectives |
| According to Ilkay, why is exploring data crucial to better modeling?<br>Data exploration... <complete the sentence><br>A. enables a description of data which allows visualization.<br>B. enables understanding of general trends, correlations, and outliers.<br>C. enables histograms and others graphs as data visualization.<br>D. leads to data understanding which allows an informed analysis of the data. | D. leads to data understanding which allows an informed analysis of the data. |
| Why is data science mainly about teamwork?<br>A. Exhibition of curiosity is required.<br>B. Data science requires a variety of expertise in different fields.<br>C. Engineering solutions are preferred.<br>D. Analytic solutions are required. | B. Data science requires a variety of expertise in different fields. |
| What are the ways to address data quality issues?<br>A. Remove outliers.<br>B. Data Wrangling<br>C. Remove data with missing values.<br>D. Generate best estimates for invalid values.<br>E. Merge duplicate records. | ACDE |

| | |
|---|---|
| What is done to the data in the preparation stage?<br>A. Cleaning, Integrating, and Packaging<br>B. Retrieve Data<br>C. Identify Data Sets and Query Data<br>D. Select Analytical Techniques<br>E. Build Models | A. Cleaning, Integrating, and Packaging |
| Which of the following is the best description of why it is important to learn about the foundations for big data?<br>A. Foundations is all that is required to show a mastery of big data concepts.<br>B. Foundations stand the test of time.<br>C. Foundations allow for the understanding of practical concepts in Hadoop.<br>D. Foundations help you revisit calculus concepts required in the understanding of big data. | C. Foundations allow for the understanding of practical concepts in Hadoop. |
| What is the benefit of a commodity cluster?<br>A. Prevents individual component failures<br>B. Prevents network connection failure<br>C. Much faster than a traditional super computer<br>D. Enables fault tolerance | D. Enables fault tolerance |
| What is a way to enable fault tolerance?<br>A. Distributed Computing<br>B. Better LAN Connection<br>C. Data-Parallel Job Restart<br>D. System Wide Restart | C. Data-Parallel Job Restart |
| What are the specific benefit(s) to a distributed file system?<br>A. High Concurrency<br>B. Data Scalability<br>C. Large Storage<br>D. High Fault Tolerance | ABD |
| Which of the following are general requirements for a programming language in order to support big data models?<br>A. Optimization of Specific Data Types<br>B. Utilize Map Reduction Methods<br>C. Enable Adding of More Racks<br>D. Support Big Data Operations<br>E. Handle Fault Tolerance | ACDE |
| What does IaaS provide?<br>A. Software On-Demand<br>B. Computing Environment<br>C. Hardware Only | C. Hardware Only |
| What does PaaS provide?<br>A. Hardware Only<br>B. Software On-Demand<br>C. Computing Environment | C. Computing Environment |
| What does SaaS provide?<br>A. Hardware Only<br>B. Computing Environment<br>C. Software On-Demand | C. Software On-Demand |
| What are the two key components of HDFS and what are they used for?<br>A. FASTA for genome sequence and Rasters for geospatial data.<br>B. NameNode for block storage and Data Node for metadata.<br>C. NameNode for metadata and DataNode for block storage. | C. NameNode for metadata and DataNode for block storage. |
| What is the job of the NameNode?<br>A. Coordinate operations and assigns tasks to Data Nodes<br>B. Listens from DataNode for block creation, deletion, and replication.<br>C. For gene sequencing calculations. | A. Coordinate operations and assigns tasks to Data Nodes |

| | |
|---|---|
| What is the order of the three steps to Map Reduce?<br>A. Map -> Reduce -> Shuffle and Sort<br>B. Shuffle and Sort -> Reduce -> Map<br>C. Map -> Shuffle and Sort -> Reduce<br>D. Shuffle and Sort -> Map -> Reduce | C. Map -> Shuffle and Sort -> Reduce |
| What is a benefit of using pre-built Hadoop images?<br>A. Quick prototyping, deploying, and validating of projects.<br>B. Less software choices to choose from.<br>C. Guaranteed hardware support.<br>D. Quick prototyping, deploying, and guaranteed bug free. | A. Quick prototyping, deploying, and validating of projects. |
| What are some examples of open-source tools built for Hadoop and what does it do?<br>A. Giraph, for SQL-like queries.<br>B. Zookeeper, analyze social graphs.<br>C. Pig, for real-time and in-memory processing of big data.<br>D. Zookeeper, management system for animal named related components. | D. Zookeeper, management system for animal named related components. |
| What is the difference between low level interfaces and high level interfaces?<br>A. Low level deals with storage and scheduling while high level deals with interactivity.<br>B. Low level deals with interactivity while high level deals with storage and scheduling. | A. Low level deals with storage and scheduling while high level deals with interactivity. |
| Which of the following are problems to look out for when integrating your project with Hadoop?<br>A. Infrastructure Replacement<br>B. Random Data Access<br>C. Task Level Parallelism<br>D. Advanced Alogrithms<br>E. Data Level Parallelism | ABCD |
| As covered in the slides, which of the following are the major goals of Hadoop?<br>A. Handle Fault Tolerance<br>B. Facilitate a Shared Environment<br>C. Enable Scalability<br>D. Optimized for a Variety of Data Types<br>E. Provide Value for Data<br>F. Latency Sensitive Tasks | ABCDE |
| What is the purpose of YARN?<br>A. Allows various applications to run on the same Hadoop cluster.<br>B. Enables large scale data across clusters.<br>C. Implementation of Map Reduce. | A. Allows various applications to run on the same Hadoop cluster. |
| What are the two main components for a data computation framework that were described in the slides?<br>A. Resource Manager and Node Manager<br>B. Resource Manager and Container<br>C. Node Manager and Applications Master<br>D. Node Manager and Container<br>E. Applications Master and Container | A. Resource Manager and Node Manager |
| (Questions 1-3 pertain to the video lecture "Exploring the Relational Data Model of CSV")<br>What is the approximate population of La Paz county in the state of Arizona for the CENSUS2010POP (column H)? (Choose the best answer.)<br>A. 25000<br>B. 10000<br>C. 15000<br>D. 20000 | D. 20000 |

| | |
|---|---|
| What county in the state of Wyoming has the smallest estimated population?<br>A. Uinta<br>B. Niobrara<br>C. Sweetwater<br>D. Platte | B. Niobrara |
| At 2:45 of the video, the Instructor creates a filter for all of the counties in California with a population greater than 1,000,000. However, included in the results is the entire state of California. This anomalous value might skew our analysis if, for example, we wanted to compute the average population of these results. What additional filter might work to resolve this problem?<br>A. Add a filter to detect and remove results which do not include the word "County" in column G.<br>B/ Add a filter which finds all counties with population greater than 100,000 AND less than 10,000,000 for column H (CENSUS2010POP).<br>C. Add a filter where the value in column E is greater than 1,000,000.<br>D. None of the above | A. Add a filter to detect and remove results which do not include the word "County" in column G. |
| (Questions 4 and 5 pertain to the video "Exploring Sensor Data")<br>How often (in seconds) do the R5 measurements occur?<br>A. 60<br>B. 50<br>C. 40<br>D. 30 | A. 60 |
| What is the field for rain accumulation?<br>A. Sm<br>B. Rc<br>C. Dx<br>D. Dn | B. Rc |
| (Questions 6 and 7 pertain to the video lecture "Exploring the Array Data Model of an Image")<br>What is the (Red, Green, Blue) pixel value for location 500, 2000?<br>A. (163, 118, 79)<br>B. (50, 156, 182)<br>C. (134, 145, 46)<br>D. (100, 123, 149) | A. (163, 118, 79) |
| Is this value likely to be land or ocean?<br>A. Land<br>B. Ocean | A. Land |
| (Questions 8 and 9 pertain to the video lecture "Exploring the Semistructured Data Model of JSON")<br>Given a tweet, what path would you most likely enter to obtain a count of the number of followers for a user?<br>A. user/followers_count<br>B. user/statuses_count<br>C. user/listed_count<br>D. None of the above | A. user/followers_count |
| Which of the following fields are nested within the 'entities' field (select all that apply)?<br>A. symbols<br>B. urls<br>C. views<br>D. tweets<br>E. user_mentions<br>F. events | ABE |
| What is a possible pitfall of utilizing Excel as a way to manipulate small databases? | |

| | |
|---|---|
| A. Excel does not allow algorithms for data manipulation.<br>B. Excel is a user program and thus cannot run on a server.<br>C. Excel does not enforce many principles of relational data models. | C. Excel does not enforce many principles of relational data models. |
| What does the term "atomic" mean in the context of relational databases?<br>A. A column or row of data. Depends on the context.<br>B. A tuple that cannot be reduced.<br>C. One unit of information that cannot be decomposed.<br>D. Fixed schema of a particular database. | C. One unit of information that cannot be decomposed. |
| What is the Pareto-Optimality problem?<br>A. Find the shortest path from source node to target node.<br>B. Find the best possible path given two or more optimization criteria where neither constraint can be fully optimized simultaneously.<br>C. Find the optimal path that requires going through specific nodes given by the user. | B. Find the best possible path given two or more optimization criteria where neither constraint can be fully optimized simultaneously. |
| What constitutes a community within a graph?<br>A. A neighborhood defined by an integer constant K around a specific node. All K+1 nodes belong in another community.<br>B. High density of nodes at a certain location.<br>C. A dense amount of edge connections between nodes in a community and a few connections across communities.<br>D. Many anomalous neighborhoods within the same vicinity | C. A dense amount of edge connections between nodes in a community and a few connections across communities. |
| Why are trees useful for semi-structured data such as XML and JSON?<br>A. They are only useful for XML data as tree-like structure is apparent with tags. While JSON does not contain a tree-like structure as it contains arrays.<br>B. Computers can easily visualize the data with a tree structure.<br>C. Trees take advantage of the parent-child relationship of the data for easy navigation.<br>D. It is not always the case that XML and JSON can be represented as trees. | C. Trees take advantage of the parent-child relationship of the data for easy navigation. |
| What is the general purpose of modeling data as vectors?<br>A. Results can be ordered by similarity using vector projection.<br>B. The ability to normalize vectors allowing probability distributions.<br>C. Enables image searching.<br>D. Enables weighting of the query. | A. Results can be ordered by similarity using vector projection. |
| For the following questions 7, 8, and 9, suppose a registration website creates data with the following fields for each person registered (note: if the user does not input a value, NULL is stored instead): Name, Date, Address, and Account Number.<br>Suppose we collect data month by month. Each month, we would have a batch of data containing the fields listed above. At the end of the year, we want to summarize our registrant activities for the entire year, so we would remove redundancies in our data by removing any records with duplicate account numbers from month to month. What type of operation do we use in this scenario?<br>A. Join<br>B. Union<br>C. Subsetting<br>D. Not an Operation | B. Union |
| From the information given in question 7, what are the constraints, if any, which we have placed on the Account Number field for the end of year collection?<br>A. There are no constraints.<br>B. Account should have at most n digits.<br>C. If we had n duplicate Account Numbers then we will remove n-1 | D. Account Number should be unique. |

| | |
|---|---|
| duplicate fields.<br>D. Account Number should be unique. | |
| Suppose 100 people signup for our system and of the 100 people, 60 of them did not input an address. The system lists the values as NULL for these empty entries in the address field. Would this situation still have structure for our data?<br>A. No because the majority of data do not have a specific field filled, thus our originally defined structure is lost.<br>B. Yes the data has structure because we have placed a structural constraint on the data, thus the data will always have the originally defined structure. | B. Yes the data has structure because we have placed a structural constraint on the data, thus the data will always have the originally defined structure. |
| What is true between data modeling and the formatting of the data?<br>A. The data does not necessarily need to be formatted in a way that represents the data model. Just so long as it can be extrapolated.<br>B. There is a one to one correspondence between formatting data and data modeling. For every model of data, there is only one way to store the data.<br>C. There is always one specific schema for storing model data that is the best and preferred method for the specific data representation. | A. The data does not necessarily need to be formatted in a way that represents the data model. Just so long as it can be extrapolated. |
| What is streaming?<br>A. Using sensors to manipulate the system, such as a smart car being able to drive by itself using sensors to detect road hazards.<br>B. Utilizing real time data to compute and change the state of an application continuously.<br>C. Calculating results using real time data otherwise known as streaming data.<br>D. Using static data stored from a real time source in order to process and guide the application. | B. Utilizing real time data to compute and change the state of an application continuously. |
| Of the following, what best describes the properties of working with streaming data?<br>A. Data is always utilized for streaming the application.<br>B. Small time windows for working with data.<br>C. Data manipulation is near real time.<br>D. Independent computations that do not rely on previous or future data.<br>E. Does not ping the source interactively for a response upon receiving the data.<br>F. Always unbounded in sequence, in other words, data is not guaranteed to be in order. | BCDE |
| What is a characteristic of streaming data?<br>A. Data is unbounded in size but requires only finite time and space to process it.<br>B. The data is unbounded in size and the size determines the time and space of processing the data.<br>C. The data is finite and requires only finite time and space to process the data.<br>D. Data is finite in size and size determines the time and space of processing the data. | A. Data is unbounded in size but requires only finite time and space to process it. |
| What type of algorithm is required for analyzing streaming data?<br>A. Fast and Complex<br>B. Accurate and Memory Efficient<br>C. Fast and Simple<br>D. Accurate and Consistent | C. Fast and Simple |
| What is lambda architecture?<br>A. A specific method for processing streaming data using special real time processes.<br>B. A specific hardware architecture for a server made specifically | |

| | |
|---|---|
| for processing real time data.<br>C. A method to process streaming data by utilizing batch processing and real time processing. | C. A method to process streaming data by utilizing batch processing and real time processing. |
| Of the following, which best represents the challenge regarding the size and frequency of data?<br>A. The size and frequency of the streaming data may be sporadic.<br>B. The size and frequency of the streaming data may be too small.<br>C. There may not be data to produce the notion of size and frequency. | A. The size and frequency of the streaming data may be sporadic. |
| What is the difference between data lakes and data warehouses?<br>A. Data lakes contain only files while data warehouses contain only databases.<br>B. Data lakes house raw data while data warehouses contain pre-formatted data.<br>C. Data lakes utilize hierarchical systems while data warehouses use object storage. | B. Data lakes house raw data while data warehouses contain pre-formatted data. |
| What is schema-on-read?<br>A. The process where formatted data is given structure when read.<br>B. Another name for data lakes.<br>C. The process where data is pre-formatted prior to being read but the schema is loaded on read.<br>D. Data is stored as raw data until it is read by an application where the application assigns structure. | D. Data is stored as raw data until it is read by an application where the application assigns structure. |
| The desired characteristics of a BDMS include (select all that apply):<br>A. Narrow range of query sizes<br>B. A full query language<br>C. Support for common "Big Data" data types<br>D. Support for ACID<br>E. Continuous data ingestion<br>F. A flexible semi-structured data model | BCEF |
| Fill in the blank with the best answer: CAP theorem states that _____ all at once within a distributed computer system?<br>A. it is necessary to have consistency, accuracy, and partial tolerance<br>B. it is impossible to have consistency, availability, and partition tolerance<br>C. it is necessary to have consistency, availability, and partition tolerance<br>D. it is impossible to have consistency, accuracy, and partial tolerance | B. it is impossible to have consistency, availability, and partition tolerance |
| What is the purpose of the acronym BASE?<br>A. To overcome CAP theorem.<br>B. The same as ACID.<br>C. Enables stricter enforcement of ACID type design.<br>D. To impose properties on a BDMS in order to guarantee certain results. | D. To impose properties on a BDMS in order to guarantee certain results. |
| What are ziplists in Redis?<br>A. A compressed list that is stored within the value of the database.<br>B. A special type of data type that can store up to 512 mb of image data.<br>C. A look up table that is stored as a value in the database. Look up table points to actual values in memory.<br>D. A special type of data type that can store hashes that point to multiple attributes. | A. A compressed list that is stored within the value of the database. |
| What is one of the main features of Aerospike?<br>A. Enables real time data streaming from external sources.<br>B. Better equipped for string based search applications. | C. Support for geospatial data storage and geospatial queries. |

| | |
|---|---|
| C. Support for geospatial data storage and geospatial queries.<br>D. Images as values within the database. | |
| What database would be best suited for the following scenario: An app development company is trying to implement a cloud based storage system for their new map-based app. The cloud will manage the longitude and latitude of the data in order to track user location.<br>A. Vertica<br>B. Redis<br>C. Aerospike<br>D. Solr | C. Aerospike |
| What database would be best suited for the following scenario: A big wholesale company is trying to implement a search engine for their products.<br>A. Redis<br>B. Solr<br>C. Vertica<br>D. Aerospike | B. Solr |
| Which of the following data types are supported by Redis? (select all that apply)<br>A. Hashes<br>B. Streaming Video<br>C. Lists<br>D. Images<br>E. Sorted Sets<br>F. Strings | ACEF |