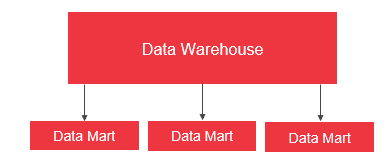
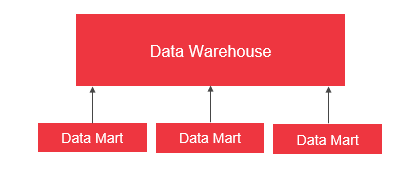
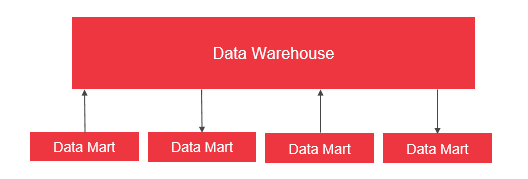
**Assessment Questions**

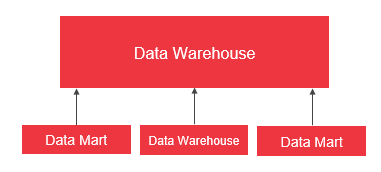
Answer the following multiple choice question (30 Question in 60 minutes)

1. Which of the following best describe the role of a data engineer?
2. Defining and enforcing rules and policies of data quality, security, governance, risk and regulatory compliance.
3. Conduct data analysis using statistical techniques to find trends and patterns in data sets to improve business
4. To assemble large complex data sets from various system to meet both functional and non-functional requirements
5. Identify opportunities within an organization and propose solution to achieve business goals
6. What do data warehouses support?
   1. OLAP
   2. OLTP
   3. OLAP and OLTP
   4. Operational Data Store (ODS)
7. Which of the following best describe operational data store (ODS)?
8. A database that is based on historical data.
9. A database that is used to run the business in real time and is based on current data.
10. A database that contain the full organization data
11. A database that primarily contain historical data
12. What is the disadvantage of normalization?
13. Pieces of a query may be distributed across many tables
14. Reduce organization of the data
15. Integrity is not maintained
16. Number of tables decreased
17. Which of the following is a top-down modelling technique for analytical system?
18. Kimball
19. Kumon
20. Pokemon
21. Inmon
22. Which of the following diagram depicts the Kimball modelling technique?









1. What is Denormalization?
2. Process of adding data redundancy in the form of extra attributes in existing table, creating new table and create an instance of an existing table
3. Process of organizing data into multiple related table to reduce data redundancy
4. Process of creating relationship between tables
5. Process of introducing uniqueness in the data
6. Which of the following is not part of NoSQL’s category?
7. Graph
8. Columnar
9. Document
10. Tabular
11. Dimensional Modelling will help users to read, aggregate and analyse numerical data in a faster and more intuitive way. (True/False)
12. Which type of fact table is most suited to store the customers’ bank account balances?
13. Transaction
14. Periodic snapshot
15. Accumulative snapshots
16. Temporal snapshot
17. Reorder the steps to implement a Type 6 Slowly Changing Dimensions (SCD).
18. Create a new column
19. Add a new record encompassing the change
20. Store the most recent past value in the new column
    1. 2 -> 1 -> 3
    2. 1 -> 2 -> 3
    3. 3 -> 2 -> 1
    4. 2 -> 3 -> 1
21. Change Data Capture (CDC) is essential because it helps to determine and track changes in data. (True / False)
22. What is the full form of SQL?
23. Simple Query List
24. Structured Query Language
25. Simple Query Language
26. None of the above
27. \_\_\_\_\_\_\_ is a collection of objects in a database.
28. Sequence
29. Schema
30. Table
31. Index
32. Select the statement correctly defining ‘View’.
33. Stores the data derived from the referenced table.
34. Does not actually store data, but derive it from tables referenced in the view definition.
35. Is the basic unit of data storage.
36. Both (b) and (c).
37. Select the commands defining Data Manipulation language(DML)
38. SELECT
39. UPDATE
40. DELETE
41. All the above
42. Which of the following SQL SELECT statement is correct?
43. SELECT col1, col2 and col3 FROM table1.
44. SELECT col1, col2, col3 FROM table1.
45. SELECT col1, col2 WHERE table1=”Correct Table”.
46. None of the above.
47. Find the names of cities with temperature more than 45 degree Fahrenheit and weather either ‘sunny’ or ‘cloudy’. (Table’s name= cityhouse)
48. SELECT city, temperature, weather FROM cityhouse WHERE weather =’sunny’ AND weather =’cloudy’ OR temperature > 45
49. SELECT city, temperature, weather FROM cityhouse WHERE weather =’sunny’ OR

weather =’cloudy’ OR temperature >45

1. SELECT city, temperature, weather FROM cityhouse WHERE weather =’sunny’ AND weather =’cloudy’ AND temperature >45
2. SELECT city, temperature, weather FROM cityhouse WHERE weather =’sunny’ OR weather =’cloudy’ AND temperature > 45
3. Below are the benefits of Alteryx tool.
4. Automate tasks
5. Increase efficiency
6. Facilitate ETL process
7. All the above
8. \_\_\_\_\_\_ returns the common records between intersecting tables, based on the common columns defined in both the tables.
9. Left Unjoin
10. Inner join
11. Right Unjoin
12. Union
13. The 3 V’s of big data are,
14. Volume
15. Velocity
16. Variety
17. Vicious
18. Data Lake follows ‘Schema-on-Read’ data format (True/False)
19. Stores data in its native form.
20. Data Warehouse
21. Data Lake
22. Database
23. None of the above.
24. HDFS is the storage component of Hadoop.(True/False)
25. Publish and subscribe streams of records, similar to a message queue or enterprise messaging system’ defines.
26. Spark
27. Kafka
28. Hive
29. Sqoop
30. Which of the following is/are example(s) of NoSQL databases.
31. MongoDB
32. Cassandra
33. RDBMS
34. Neo4j
35. \_\_\_\_\_\_\_\_ stores data in the form of vertices and edges.
36. Graph NoSQL Database
37. Columnar NoSQL Database
38. Key-Value NoSQL Database
39. Document NoSQL Database
40. Document NoSQL stores values in the form of \_\_\_\_\_\_.
41. Key and Value
42. JSON and BSON
43. Columnar structures
44. None of the above
45. What is “C” in CAP Theorem?
46. Common
47. Complex
48. Consistency
49. Capital
50. What ETL Stands for
51. Engage, Transform, Load
52. Extract, Travel, Load
53. Extract, Transform, Load
54. Extract, Transform, Landing