1.- A que se refiere el termino KDD

A) Knowledge Discovery in Databases o descubrir conocimiento en las bases de datos

B) Key Database Doom

C) Keep Discovery Drag

D) Know Database

2.- Es una infraestructura digital de desarrollo creado en código abierto bajo licencia Apache.

a) ITIL

b) SAP

c) ERP

d) hadoop

3.- Es una base de datos no relacional distribuida y basada en un modelo de almacenamiento de <clave-valor>, desarrollado en Java. Permite grandes volúmenes de datos en forma distribuida.

a) Oracle

b) Cassandra

c) PG Admin

d) Docker

4- Diseñado para la colección y análisis a gran escala de “logs”. Incluyeun toolkit para desplegar los resultados del análisis y Monitoreo.

A) Big data

B) Command

C) Chukwa

D) File over

5.- El siguiente es un ejemplo en HQL para crear una tabla, cargar datos yobtener información de la tabla utilizando???

a) Sql

b) Oracle

c) Hive

d) Msdos

6.- Su tarea principal es dirigir los datos de una fuente hacia alguna otra localidad, en este caso hacia el ambiente de Hadoop.

a) inpath

b) No Sql

c) flume

d) java

7.- Una plataforma de alto nivel para el procesamiento de datos de gran tamaño en clusters de Hadoop.

a) hdi

b) EBP

c) PIG

d) IBM

8.- Transfiere grandes cantidades de datos entre Hadoop y bases de datos relaciónales, como SQL u otros almacenes de datos estructurados, de la forma más eficiente posible.

a) sqoop

b) PBI

c) know how

d) Hadoop

9.- Datos ordenados y bien definidos en cuanto a su formato, tamaño y longitud. De esta manera, es más fácil procesarlos. Por ejemplo, base de datos, hojas de cálculo, etc.

a) datos falsos

b) datos semiestructurados

c) datos numéricos

d) datos estructurados

10.- Lugar donde se almacenan los datos con fines de archivamiento, análisis de efectos y seguridad.

a) BD CONTINUA

b) nube

c) data warehouse

d) servidor

11.- Data Warehouse en \_\_\_\_\_\_ son almacenes de datos en los que se actualiza cada momento a medida que llegan nuevos datos.

a) decimal

b) fura de línea

c) simétricos

d) tiempo real

12.- Cual no es una característica de un Data Warehouse.

a) simétricos

b) Integrado

c) Variante de tiempo

d) Orientado a Sujetos

13.- Es el proceso de alterar los datos en un almacenamiento para asegurarse de que son exactos y correctos.

a) datos alternos

b) limpieza de datos

c) datos simétricos

d) datos almacenados

14.- Permite obtener conectividad entre SAS y Hadoop, vía Hive.

a) SAS/ACCESS Interface to Hadoop

b) <sas access>

c) <out sas access>

d) Software SAS Base

15.- Realiza exploración visual de los datos para descubrir nuevos patrones y publicar informes en la web y dispositivos móviles.

a) SAS In-Memory Statistics for Hadoop

b) SAS Visual Statistics

c) SAS/ACCESS

d) SAS Visual Analytics

16.- Es una solución líder en la industria construida sobre una plataforma de calidad de datos que le ayudará a mejorar, integrar y gobernar sus datos.

a) HOOP

b) SAS DATA MANAGEMENT

c) BIG DATA

d) RENDER