

# Teknologi Perekayasaan Data Quiz 2

222011405@stis.ac.id [Switch account](#)



Draft saved

Your email will be recorded when you submit this form

**Silakan pilih salah satu jawaban yang paling tepat!**

Dari beberapa format file di bawah ini, manakah yang digunakan untuk menyimpan data terstruktur dalam Hadoop?

- ☒ CSV File Format ✓
- ☐ Sequence File Format
- ☐ Avro File Format ✓
- ☐ JSON File Format ✓

[Clear selection](#)

Sqoop menggunakan \_\_\_\_\_ untuk mengambil data dari RDBMS dan menyimpannya dalam HDFS

- ☐ Hive
- ☐ HBase
- ☒ MapReduce
- ☐ Pig



Clear selection

Tujuan dari perintah **sqoop import** di Sqoop Hadoop adalah

- ☐ Mengimport data dari Hadoop ke database relasional
- ☒ Mengimpor data dari database relasional ke Hadoop
- ☐ Mengimpor data dari satu tabel Hadoop ke tabel lainnya di Hadoop
- ☐ Mengimpor data dari satu database relasional ke tabel lainnya di basis data yang sama



Clear selection

Jika diketahui suatu tabel Hive **student** berisi 100 baris data tersedia di Hadoop. Tabel-tabel lain dibuat dengan perintah di bawah ini dan diisi dengan data yang sama dengan **student**.

Manakah dari table-tabel tersebut yang memiliki file size paling kecil?

```
CREATE TABLE IF NOT EXISTS table1 LIKE test STORED AS PARQUET;  
CREATE TABLE IF NOT EXISTS table2 LIKE test STORED AS TEXTFILE;  
CREATE TABLE IF NOT EXISTS table3 LIKE test STORED AS ORC;
```

- ☐ table1
- ☐ table2
- ☒ table3



Clear selection

Dalam Apache Hive, apa perbedaan antara Managed Table dan External Table?

- ☐ Managed Table adalah tabel yang dapat diakses secara eksternal oleh aplikasi lain di luar Apache Hive, sedangkan External Table adalah tabel yang hanya dapat diakses oleh Apache Hive itu sendiri
- ☐ Managed Table adalah tabel yang memerlukan konfigurasi khusus untuk mengakses data eksternal, sedangkan External Table adalah tabel yang data fisiknya disimpan dalam basis data Hive secara terorganisir
- ☐ Managed Table dan External Table adalah istilah yang digunakan secara bergantian untuk menggambarkan tabel-tabel yang dikelola oleh Apache Hive, tidak ada perbedaan fungsional antara keduanya
- ☒ Managed Table adalah tabel yang disimpan di dalam basis data Hive secara terorganisir, sedangkan External Table adalah tabel yang hanya menyimpan metadata tanpa menyimpan data fisik

Clear selection

Pernyataan berikut yang benar adalah

- ☐ HBase memproses data dalam Hadoop menggunakan MapReduce
- ☐ HBase mengimplementasikan model data document seperti halnya database NoSQL lain
- ☒ Random access pada HBase memungkinkan karena adanya rowkey pada setiap baris data
- ☐ Pada HBase, data di Region dibagi-bagi berdasarkan rentang column untuk mendistribusikan beban secara merata

Clear selection

Perintah HBase berikut digunakan untuk membaca baris/cell data adalah

- ☐ select
- ☒ get
- ☐ put
- ☐ fetch

Clear selection

Jika kita melakukan eksekusi Pig script, komponen Parser akan memeriksa sintaks script dan menghasilkan

- ☒ Logical plan ✓
- ☐ Optimized logical plan
- ☐ Physical plan
- ☐ MapReduce jobs

Clear selection

Suatu data berisi 10 baris data dengan schema sebagai berikut. Apakah output dari Pig script tersebut?

```
data = LOAD '/input/data.csv' USING PigStorage(',') AS (id: int, name: chararray, age: int);
filtered_data = FILTER data BY age > 25;
grouped_data = GROUP filtered_data BY name;
result = FOREACH grouped_data GENERATE group, COUNT(filtered_data);
STORE result INTO '/output/result';
```

- ☐ Jumlah entri untuk setiap kelompok usia dan nama dengan usia lebih dari 25
- ☐ Jumlah entri untuk setiap kelompok nama
- ☐ Jumlah entri untuk usia lebih dari 25
- ☒ Jumlah entri untuk setiap kelompok nama dengan usia lebih dari 25 ✓

Clear selection

Apache Kafka menggunakan model publish-subscribe untuk \_\_\_\_

- ☐ Mengirim pesan ke satu konsumen secara langsung
- ☒ Mengirim pesan ke beberapa konsumen secara langsung ✓
- ☐ Mengirim pesan ke satu konsumen secara terjadwal
- ☐ Mengirim pesan ke beberapa konsumen secara terjadwal

Clear selection



Apakah perintah untuk membuat topic pada Apache Kafka?

- ☐ kafka-topics.sh ✓
- ☐ kafka-cli.sh
- ☐ kafka-producer.sh
- ☒ kafka-create-topic.sh

Clear selection

Apa manfaat utama dari lazy evaluation di Apache Spark?

- ☐ Mengurangi latensi waktu respon dalam pemrosesan data
- ☒ Mengurangi penggunaan memori yang berlebihan
- ☐ Meningkatkan keamanan data yang diproses
- ☐ Meningkatkan keakuratan hasil analisis data

Clear selection

Yang merupakan Transformasi RDD dalam Apache Spark adalah

- ☐ Operasi yang menggabungkan dua RDD menjadi satu
- ☒ Operasi yang mengubah konten RDD dengan mengaplikasikan fungsi
- ☐ Operasi yang mengurangi RDD menjadi elemen tunggal
- ☐ Operasi yang mengubah RDD menjadi tipe data yang berbeda

Clear selection



Apa yang dimaksud dengan ekstraksi fitur (feature extraction) dalam pemrosesan unstructured data?

- ☐ Proses mengubah data tidak terstruktur menjadi format tabel terstruktur
- ☐ Proses menghilangkan data yang tidak relevan dari data tidak terstruktur
- ☒ Proses mengubah data tidak terstruktur menjadi data terstruktur dengan mengidentifikasi fitur penting
- ☐ Proses menggabungkan beberapa file data tidak terstruktur menjadi satu

Clear selection

Apa yang dimaksud dengan schema on read dalam pemrosesan data semi terstruktur di Hadoop?

- ☐ Penentuan struktur data saat data ditulis ke HDFS
- ☒ Penentuan struktur data saat data dibaca atau diproses dari HDFS
- ☐ Proses transformasi data semi terstruktur menjadi format terstruktur sebelum diproses
- ☐ Proses mengubah data semi terstruktur menjadi data terstruktur dengan skema yang telah ditentukan

Clear selection

[Back](#)

[Submit](#)

[Clear form](#)

Never submit passwords through Google Forms.

This form was created inside of Politeknik Statistika STIS.

[Report Abuse](#)

Google Forms