**NAMA** **:** RIZKI ARIF SETIADI

**NIM :** 145410154

**Soal UTS Pengenalan Big Data**

1. Cari dan sebutkan 3 DBMS yang bisa digunakan untuk mengelola big data

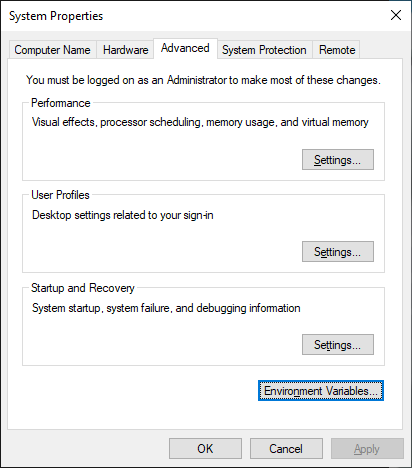
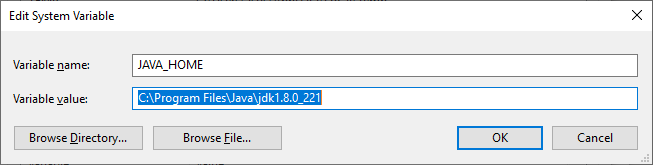
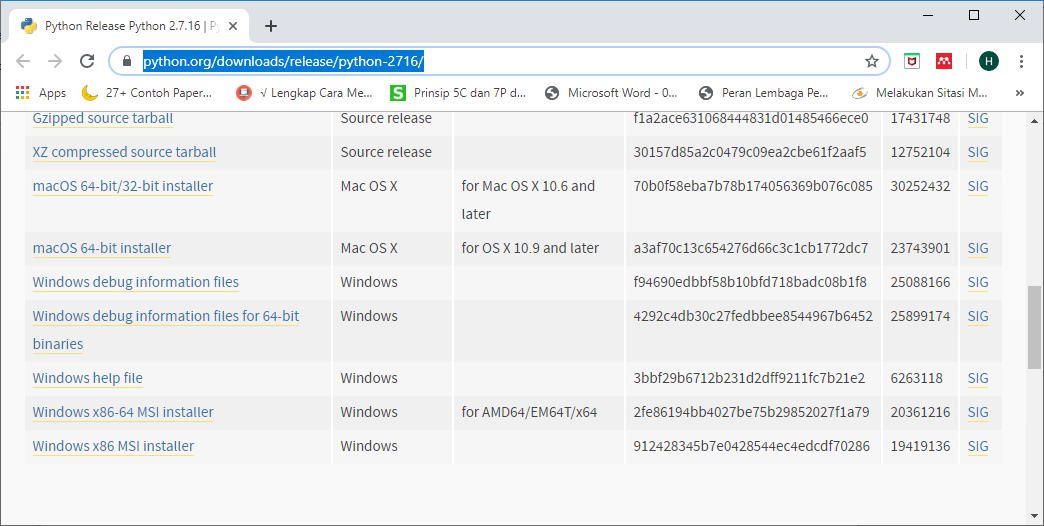
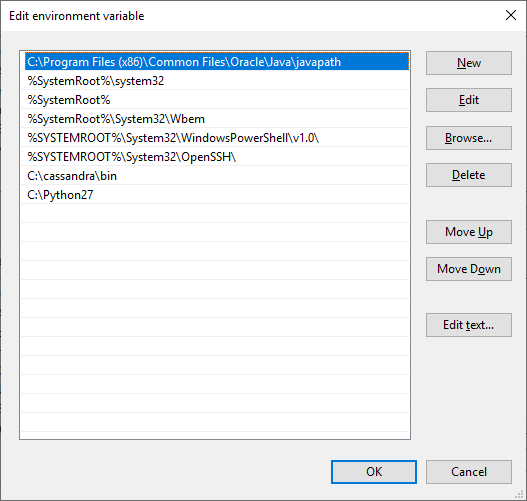
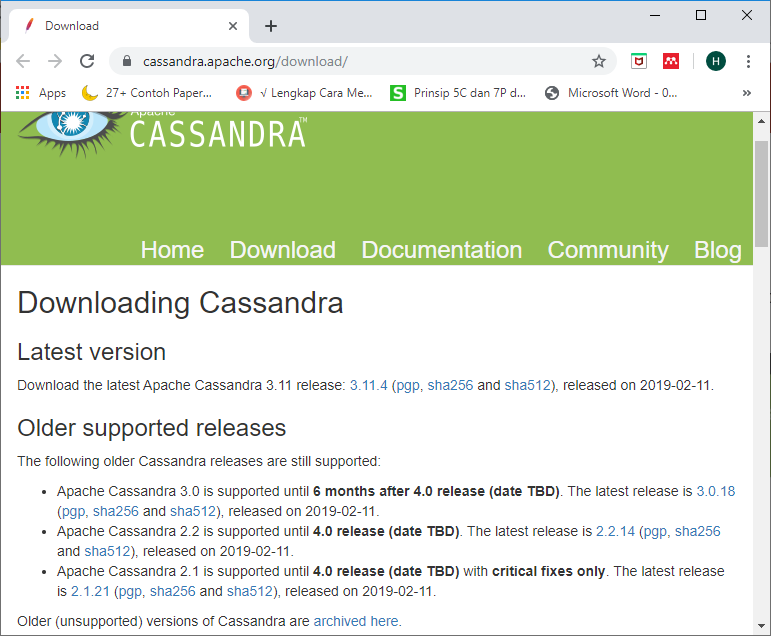
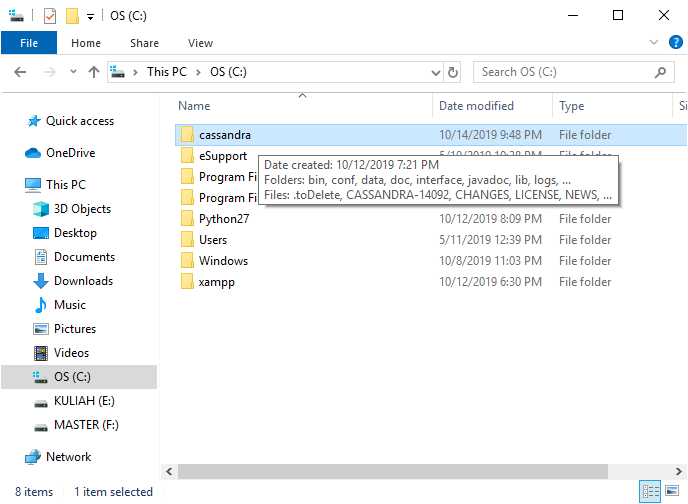
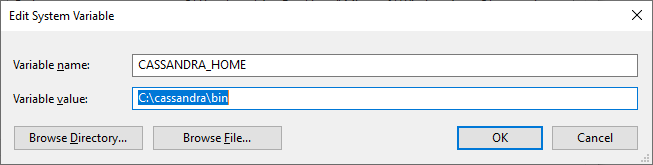
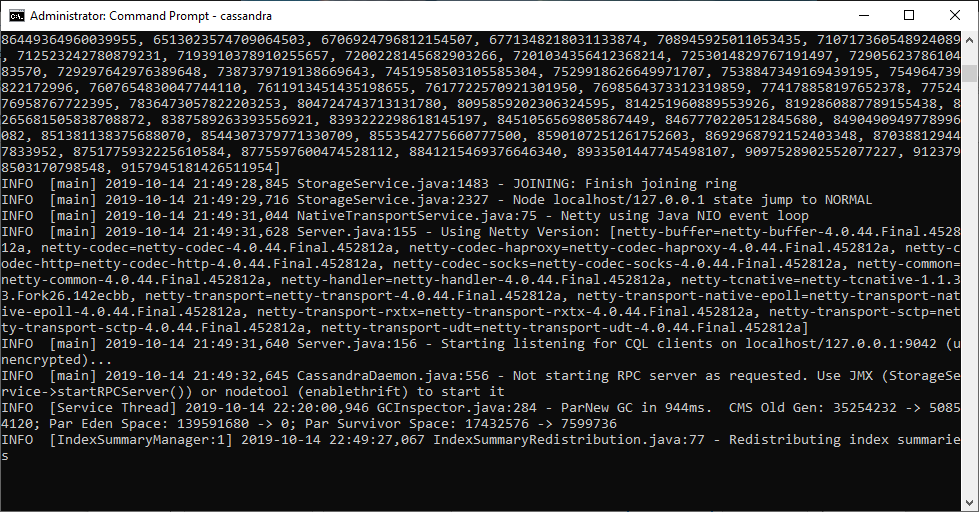
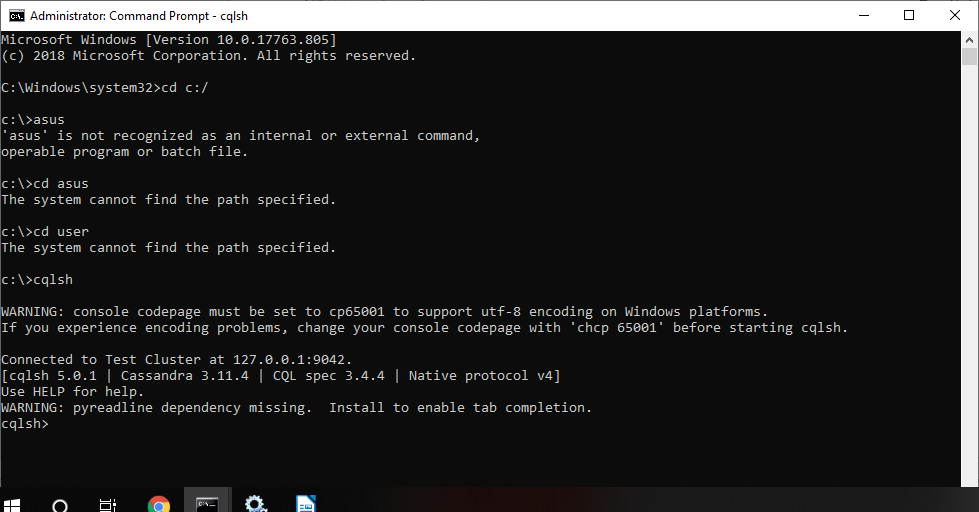
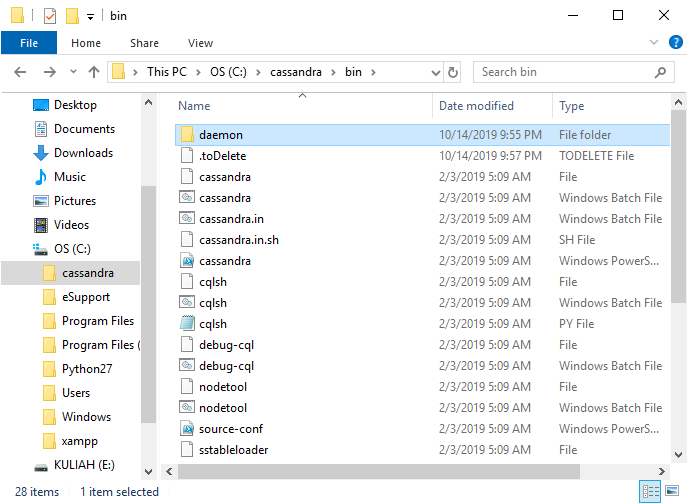
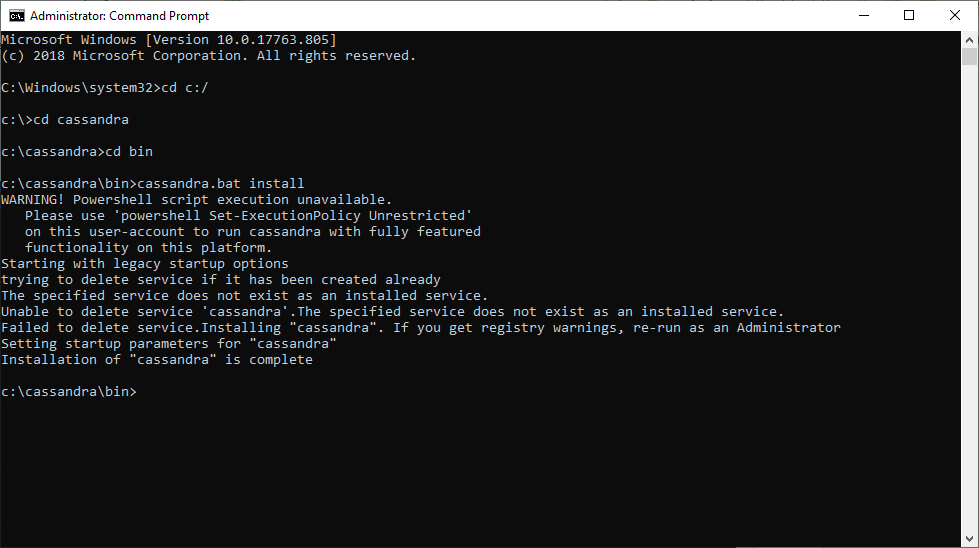
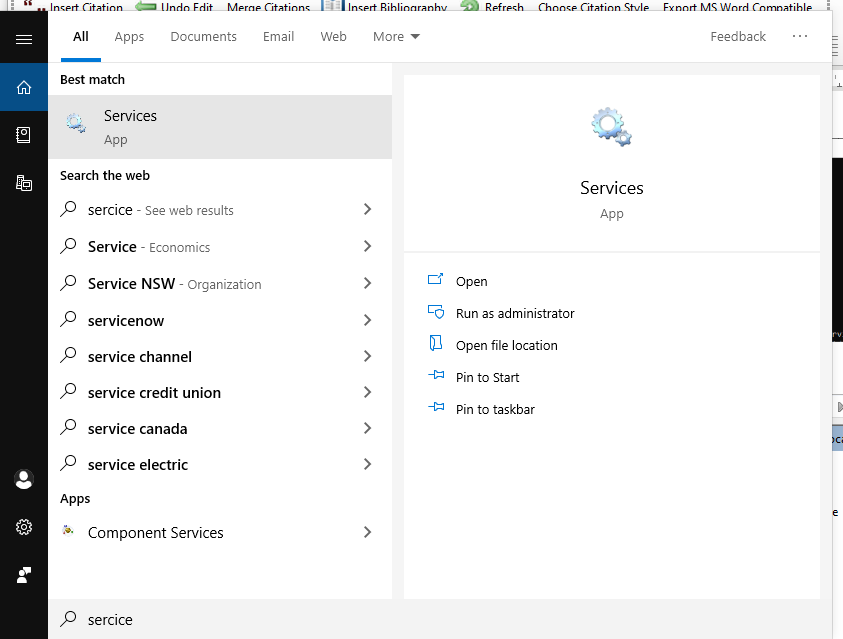
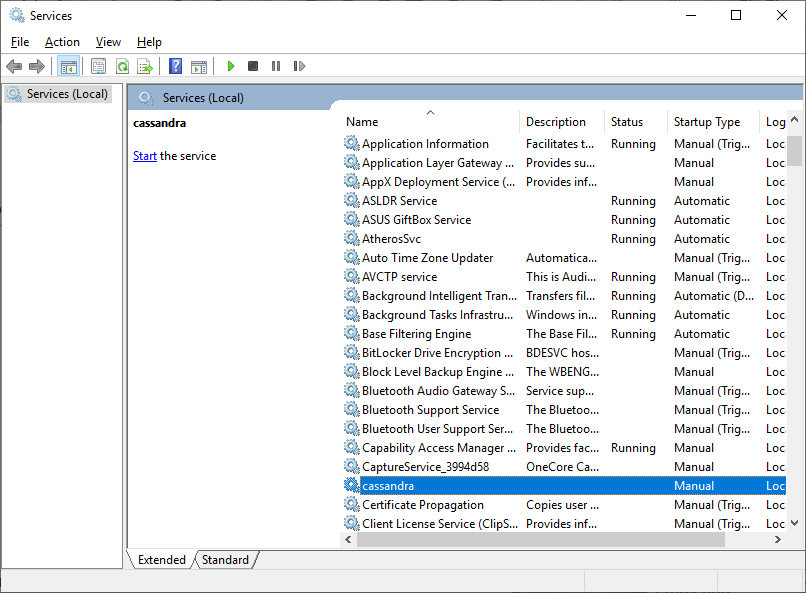
2. Carilah contoh masalah big data yang bisa dikelola menggunakan salah satu DBMS tersebut, jelaskan mulai dari instalasi sampai CRUD untuk data menggunakan DBMS tersebut. Asumsikan anda akan memecahkan masalah big data yang sudah anda cari contoh tadi, jelaskan kira-kira bagaimana arsitektur dari solusi big data menggunakan DBMS tersebut, gambarkan diagramnya.

**Jawaban :**

1. -Cassandra

-MongoDB

-Hive

* + Langkah pertama, download java 8 ( Java SE development Kit 8u211) disini https://www.oracle.com/technetwork/java/javase/downloads/jdk8-downloads-2133151.html
  + pilih yang sesuai sistem operasi kita yaitu windows 64bit.
  + Setelah kita download, jalankan instalasi dan kemudian kita lanjutkan ke langkah setting variabel JAVA\_HOME di windows kita seperti gambar dibawah ini :
  + 
  + klik pada menu environment variables, klik new pada system variables dan masukkan variable name dengan “ JAVA\_HOME” lalu isi variable value dengan klik browse directory, arahkan ke lokasi penginstalan java sdk yaitu di C:\Program Files\Java\jdk1.8.0\_221 lalu klik OK dan OK lagi. Mari kita lanjutkan tahap berikutnya yaitu mendownload dan menginstall python 2.7.
  + 
  + Kamu bisa mendownload python versi tersebut disini : https://www.python.org/downloads/release/python-2716/ dan pilih sesuai windows kita yaitu 64bit.
  + 
  + Setelah itu install python 2.7 dan selanjutnya kita set system environment variable untuk python 2.7. kembali ke cara yang sebelumnya buka system environment variable yang ada di windows lalu pilih system variable dan pilih pada variable path, nanti akan muncul menu edit environment variable, pilih saja new dan browse, arahkan ke directori python yang kita instal yaitu di C:\Python27. Tampilannya seperti ini :
  + 
  + next, Kita lanjutkan dengan mendownload apache cassandra yang ada pada web ini https://cassandra.apache.org/download/ dan ambil versi terbarunya.
  + 
  + Ok setelah didownload silakan extract file yang kita download tadi di directory C:// dan rename folder tersebut menjadi “cassandra” untuk mempermudah kita memanggil aplikasi di command prompt.
  + 
  + Selanjutnya kita lanjut lagi setting CASSANDRA\_HOME untuk nama variabel nya dan directori nya arahkan ke C:\cassandra\bin , tampilannya seperti ini :
  + 
  + lalu sama seperti langkah sebelumya, dibuat juga path pada system variable dengan mengklik path pada system variabel. Klik new dan browse directory yang sama dengan yang kita buat saat membuat variabel name cassandra\_home yaitu di C:\cassandra\bin.
  + sekarang di command prompt dan jalankan sebagai administrator langsung saja panggil dengan command “cassandra” lalu akan muncul tulisan seperti ini :
  + 
  + apabila sudah ada keterangan node localhost/127.0.0.1 state jump to normal tandanya sistem cassandra sudah masuk ke dalam windows kita. Buka lagi command prompt 1 lagi dan ketik perintah cqlsh, namun sebelumnya harus keluar dari directory default dan arahkan ke directory C:// seperti ini :
  + 
  + Sekarang cassandra kita sudah berhasil tersambung, dan bisa dimulai untuk coding. Namun langkah ini belum selesai karena untuk menggunakan cassandra kita harus menggunakan langkah seperti memanggil cassandra di command prompt. So, kita bisa setting cassandra ini sebagai services di windows kita, yaitu dengan cara :
  + 1. download apache commons daemon disini : https://archive.apache.org/dist/commons/daemon/binaries/windows/?C=M;O=D
  + pilih yang terbaru atau latest update.
  + 2. buka file .zip tersebut dan extract lalu buka file yang sudah di extract, arahkan langsung ke folder amd64 dan ambil file prunsrv.exe (untuk 32bit bisa menggunakan file prunsrv.exe yang ada diluar folder amd64). Copy file exe tersebut dan masukkan ke directory cassandra dan di dalam folder bin, buat folder baru bernama “daemon” dan paste di dalam folder “daemon” tersebut, seperti ini :
  + 
  + setelah itu kita jalankan command prompt lagi menggunakan perintah seperti dibawah ini :
  + 
  + selanjutnya mari kita set services cassandra ini di windows dan jalankan sebagai administrator :
  + 
  + lanjut kita cari services nya cassandra disana lalu start the services. kini cassandra menjadi services yang aktif di windows kita :
  + 

**Arsitektur :**

* Arsitektur
* Cassandra di desain awal untuk menghandle Big Data yang terdiri dari banyak titik-titik (node) yang terpisah-pisah dan saling bekerjasama nyaris tanpa ada kesalahan.
* Cassandra memiliki peer-to-peer sistem terdistribusi di seluruh node, dan data didistribusikan di antara semua node dalam sebuah cluster.
* Semua node dalam sebuah cluster memainkan peran yang sama. Setiap node independen dan pada saat yang sama saling berhubungan ke node lain.
* Setiap node dalam sebuah cluster dapat menerima membaca dan menulis permintaan, terlepas dari mana data sebenarnya terletak di cluster.
* Ketika sebuah node performanya turun, membaca permintaan / tulis dapat dilayani dari node lain dalam jaringan.
* Replikasi data di Cassandra disebut dengan istilah Gossip Protocol dimana satu atau lebih node dalam sebuah Cluster sebagai replika untuk bagian tertentu dari data. Jika terdeteksi bahwa beberapa node datanya out of date, Cassandra akan mengembalikan nilai terbaru untuk klien. Setelah mendapatkan nilai kembalian terbaru, Cassandra melakukan perbaikan membaca di latar belakang untuk memperbarui nilai-nilai yang out of date.

**Diagram :**

