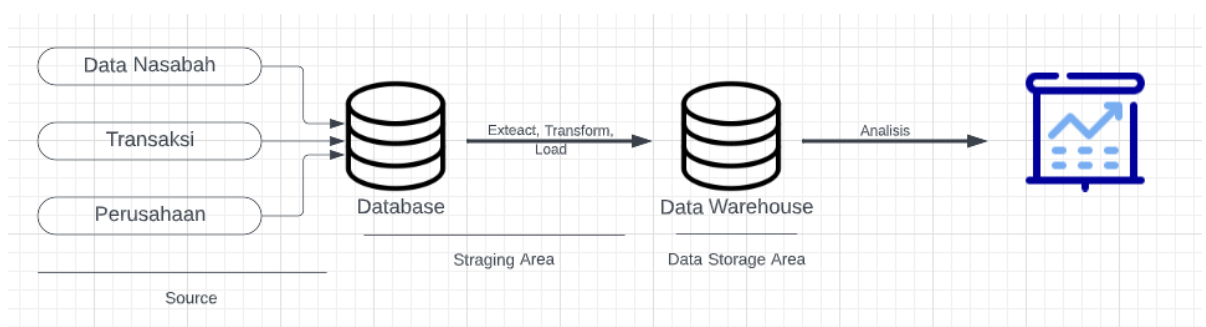


Nama Kursus : Microsoft Azure Data Fundamentals
Pertemuan : Asinkron 6
Tim Penyusun : Charles Bernando, S.Si, M.A., Ph.D.
Mentee : Amelia Angraini M
Jalur : Data Analyst
Total Jam Pembelajaran : 25 JP

Untuk kasus di dunia perbankan:

1. Buatlah sebuah diagram yang menjelaskan alur data berjalan dari database hingga ke data warehouse sampai akhirnya digunakan untuk visualisasi.

Jawaban:



2. Pada kasus dimana diperlukan data social media seperti data dari Twitter, komponen Azure apa yang dapat digunakan untuk memasukkan data social media tersebut ke dalam data warehouse? Jelaskan juga proses dan tahapan masuknya data dari social media ke data warehouse!

Jawaban:

Dalam pembahasan kasus, Azure Stream dapat digunakan sebagai alat Analisa untuk membaca serta menganalisis data media sosial seperti Twitter, hingga sosial media lainnya. Dalam analisis, sosial media tergolong sebagai data stream yang dapat dijadikan bahan tinjauan dan diproses dan dianalisis secara *real time* Ketika sirkulasi data baru terjadi.

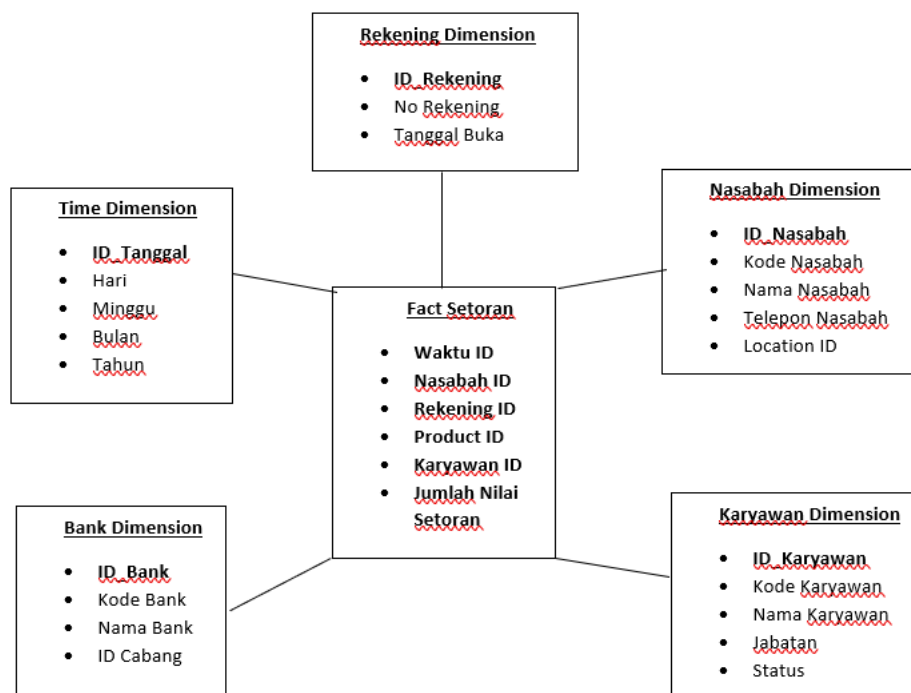
Untuk menyajikan data sosial media tersebut ke dalam *data warehouse*, komponen-komponen sebagai berikut dapat digunakan, antara lain **Azure Stream Analytics**, **Azure Data Explorer**, **Spark Structured Stream** dimana komponen-komponen tersebut dapat dikaji dan digunakan dalam stream processing. Selain itu, hal-hal sebagai berikut juga menjadi kewajiban komponen agar pemrosesan data dapat dilakukan;

- Memiliki aplikasi *TwitterClientCore* untuk membaca dan menganalisis umpan atau konten yang ada pada Twitter
- Memiliki akun *Azure*
- Memiliki akun *Twitter* dan komponen kajian terkait
- Menginstall *.NET Core* versi 2.1.0

Berikut adalah proses serta tahapan yang dapat dilakukan dalam memecahkan kasus:

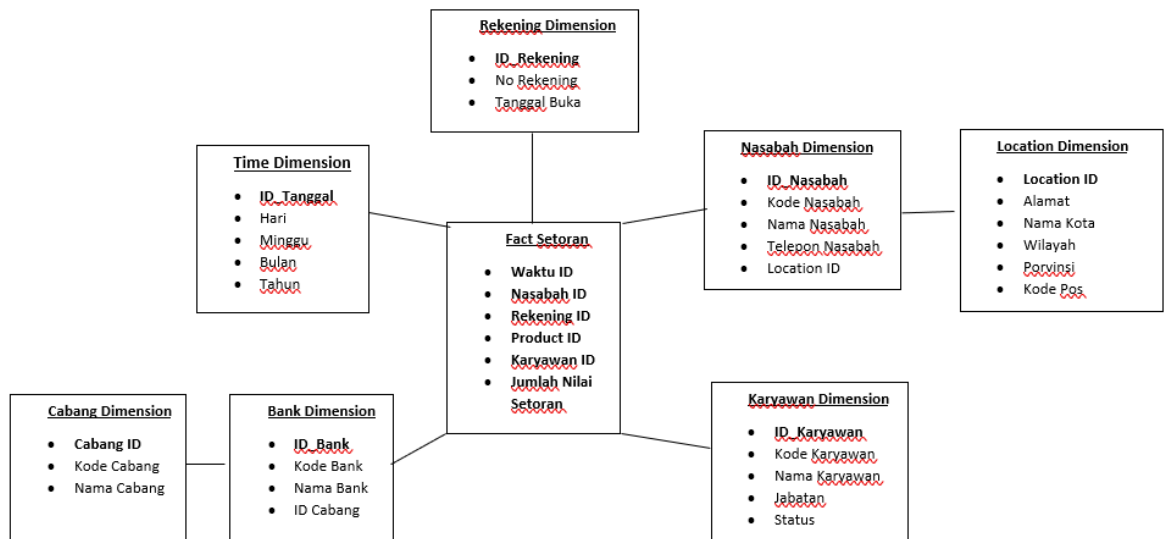
1. Membuat hub kumpulan kejadian untuk *input streaming* yang kemudian digunakan dalam mengelompokkan secara instan kejadian terkait secara logis dan sesuai formula
 2. Data yang telah dihasilkan kemudian direkam oleh sumber streaming seperti *Azure Event Hubs*, *Azure IoT Hub*, *Azure Data Lake Store Gen 2*, dan *Apache Kafka* untuk diproses melalui *cloud* ataupun *table database*.
 3. Data yang telah dihasilkan dan direkam lalu diproses dan dilakukan pengkuerian atau juga melalui prosedur ETL, dan selanjutnya data akan diolah kedalam data warehouse.
 4. Hasil dari olahan pemrosesan ditulis kedalam *output* (atau *sink*) dalam bentuk berupa file, tabel database, *dashboard visual* secara *real time*, maupun *queue* lain untuk diproses lebih lanjut dengan *query* berikutnya.
3. Buatlah sebuah skema data warehouse dengan menggunakan star schema dimana terdapat satu fact table dan beberapa dimension tables. Berikan juga atribut yang ada pada fact table dan dimension tablesnya (atribut dibebaskan)

Jawaban:



4. Buatlah sebuah skema data warehouse dari skema di nomor 3 diatas, namun kali ini dengan menggunakan skema snowflake. Kemudian berikan kelebihan dan kekurangan skema snowflake ini dibandingkan dengan star schema di nomor 3!

Jawaban:



Kelebihan dan kekurangan skema snowflake ini dibandingkan dengan snowflake yaitu:

Star Schema		Snowflake Schema	
Kelebihan	Kekurangan	Kelebihan	Kekurangan
Cenderung mudah dipahami karena modelnya yang lebih sederhana	Ukuran data lebih besar dikarenakan data yang disimpan ulang	Ukuran data lebih kecil di dalam data penyimpanan	Cenderung lebih sulit dipahami dikarenakan kompleksitasnya
Memudahkan mencari isi karena kesederhanaannya dengan cara melihat step by step dari masing-masing dimensinya	Maintenance dan update lebih sulit	Lebih mudah dilakukan maintenance dan update	Sulit mencari data yang dibutuhkan karena melihat strukturnya yang kompleks dan bercabang-cabang
Proses query yang lebih singkat		Proses query lebih cepat pada saat proses ETL	