

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Data adalah sebuah kumpulan fakta, konsep yang kemudian diolah menjadi sebuah informasi yang dapat dimengerti dan dipahami oleh manusia (Yuliana, 2021). Data juga dikatakan sebagai bahan mentah dari informasi yang dihasilkan (Puspaningrum, 2020). Selain itu, Anantama pada tahun 2020 mendefinisikan data sebagai kenyataan yang menggambarkan sebuah kejadian nyata, yang dapat diperoleh melalui karakter, huruf, angka, simbol, suara, dan sebagainya. Agar sebuah data menjadi bermanfaat perlu pengolahan data menjadi sebuah informasi yang berguna bagi manusia (Pratama, 2021). Pengolahan data menjadi sebuah informasi yang baik dapat memberikan kesan bahwa pesan dalam informasi tersebut meyakinkan dan menciptakan pertukaran informasi yang efektif dan efisien (Mangadar, 2020).

Sebelum diolah menjadi sebuah informasi, biasanya data disimpan dalam sebuah media penyimpanan tertentu seperti basis data atau *database*. *Database* merupakan sebuah tempat untuk menyimpan kumpulan data yang saling berhubungan untuk menciptakan informasi yang lebih berguna bagi suatu organisasi atau perusahaan (Selamet Samsugi, Yusuf, 2020). Andaru (2018), basis data atau *database* merupakan kumpulan dari banyak informasi yang tersimpan secara sistematis sehingga dapat dihubungkan dan dikelola oleh program komputer. Database terdiri atas susunan dan kumpulan data yang tersimpan (Yanti, 2018).

Berdasarkan penjelasan tersebut, sebuah data yang tersimpan di dalam basis data sebagai media penyimpanan saling terhubung dan berkualitas untuk menciptakan informasi yang berguna dan bermanfaat. Akan tetapi berdasarkan pengalaman saat bekerja di PT Dua Empat Tujuh, data yang tersimpan dalam sebuah basis data tidak semuanya memiliki kualitas yang baik. Kegagalan proses pengambilan data mungkin saja terjadi sehingga pada proses load data terdapat anomali data pada basis data.

Anomali merupakan sebuah kejadian yang tidak dapat diperkirakan dan menjadi sebuah penyimpangan atau keanehan pada sebuah proses yang sedang berjalan. (Dewi, 2017). Anomali data dapat dikatakan sebagai keanehan pada sebuah data yang dihasilkan dari sebuah proses yang memberikan penyimpangan pada informasi yang dihasilkan. Anomali data ini dapat berpengaruh pada proses bisnis karena dapat menyebabkan kerugian (Amara, dkk, 2013). Untuk itu anomali data ini perlu menjadi perhatian utama agar menjadi himbauan bagi pemilik proses untuk menentukan tindakan penanganan secepat mungkin. Berdasarkan pengalaman penyusun ketika bekerja di PT Dua Empat Tujuh, anomali data seperti keterlambatan data atau tidak ketersediaan data yang terjadi belum banyak sistem informasi yang dapat memberikan pemberitahuan. Untuk itu, salah satu solusi yang mungkin dibangun adalah sebuah sistem informasi berupa sistem monitoring yang dapat digunakan sebagai himbauan bagi pemilik proses data maupun tim operasional.

Sistem monitoring sendiri merupakan seperangkat proses yang dibangun untuk melakukan penilaian atau pemantauan yang dilakukan secara terus

menerus terhadap sebuah proses untuk mendapatkan harapan atau tujuan yang telah direncanakan (Maudy dan Hari P, 2018). Monitoring ini bisa dilakukan dengan memanfaatkan beberapa tools seperti *Datapad*, *Grafana*, *Databox*, *Metabase*, dan sebagainya. *Grafana Monitoring Dashboard* dipilih dalam pembangunan sistem ini selain karena karena *opensource*, *Grafana* juga memiliki beberapa keunggulan dan keuntungan yang dapat diterima oleh pengguna. Seperti, *Grafana* memiliki opsi visualisasi data bawaan mulai dari *heatmaps*, *graph*, *geomaps* yang dapat dimanfaatkan untuk visualisasi data yang lebih cepat dan flexible. Selain itu *Grafana* juga memiliki fitur untuk mengambil sumber data melalui 40 lebih plugin dan API. Selain itu, salah satu fitur yang akan digunakan dalam pembangunan sistem ini adalah *Grafana Alert* yang dapat dimanfaatkan dalam pembuatan pelaporan yang *realtime* (John, 2023).

Sebelum menyajikan data yang dapat di monitoring oleh *Grafana*, diperlukan proses pengolahan data. Proses pengolahan data ini dapat dilakukan dengan konsep *Extract*, *Transfrom*, dan *Load* (ETL). ETL sendiri merupakan sistem atau kumpulan proses penyiapan data dari sumber data masuk sampai pada proses penyimpanan data (Yetli dan Harianto Kristanto, 2019). Dalam membangun proses ETL juga terdapat beragam tools *opensource* yang dapat digunakan, seperti *Hevo Data*, *Pentaho*, *Apache Nifi*, *Apache Airflow*, dan sebagainya. Dalam pembangunan sistem ini, *Apache Nifi* dipilih karena merupakan salah satu *tools* ETL yang dirancang untuk membantu membantu pengelolaan data dari berbagai sumber yang bersifat *realtime* (Ibnu,

2023). Kemudian, *Apache Nifi* menyediakan *user interface* berbasis web yang dapat memudahkan dalam konfigurasi yang akan dilakukan (Aakash, 2023). Berdasarkan beberapa keunggulan dari *Grafana* dan *Apache Nifi* tersebut, penyusun menggabungkan keduanya untuk membangun sistem monitoring ini.

## **1.2 Rumusan Masalah**

Sebagaimana telah diuraikan pada latar belakang diatas, penelitian ini dilakukan untuk menjawab bagaimana implementasi ETL dan visualisasi dashboard dengan *Apache Nifi* dan *Grafana Dashboard* menjadi salah satu opsi sistem monitoring anomali data.

## **1.3 Ruang Lingkup**

Dalam pembangunan sistem monitoring ini, agar lebih terarah dan fokus pada tujuannya, maka akan dibatasi oleh beberapa ruang lingkup sebagai berikut :

1. Penentuan data dapat dikatakan anomali pada sistem monitoring ini didasarkan oleh dua faktor. Pertama, ketika suatu data yang dihasilkan dari proses ETL yang tersimpan dalam database MySQL memiliki jumlah record 0. Kedua, apabila jumlah data hasil proses ETL yang tersimpan pada MySQL memiliki total revenue 0.
2. Data yang digunakan pada pembangunan sistem monitoring ini menggunakan data sample revenue supermarket yang dipublikasi oleh Itman Arif Hariri pada website *Kaggle* (2022). Dari sample data tersebut

akan diformat dengan beberapa bentuk yakni flat file, *UDP List*, dan *HTTP*. Hal ini bertujuan untuk menjadi simulator ketika input data dari format yang berbeda beda.

3. *Alert* yang dihasilkan oleh visualisasi pada *Grafana* akan dikirimkan melalui telegram sebagai media informasi. Kemudian, data *alert* tersebut tetap tersimpan pada *Grafana Dashboard*.

#### **1.4 Tujuan Penelitian**

Tujuan dari penelitian implementasi ETL dan visualisasi data dalam sistem monitoring anomali data adalah diharapkan dapat memberikan salah satu opsi penyedia informasi berupa pemberitahuan tentang anomali data yang terjadi.

#### **1.5 Manfaat Penelitian**

Diharapkan dengan dilakukannya pembangunan sistem monitoring anomali data dengan *Apache Nifi* dan *Grafana Dashboard* ini dapat memberikan salah satu opsi sistem informasi yang dapat memberikan pemberitahuan mengenai anomali data yang mungkin terjadi. Selain itu diharapkan dengan adanya penelitian ini dapat dikembangkan menjadi objek yang lebih luas atau pun pembangunan sistem dengan tools berbeda lainnya. Kemudian diharapkan dengan dilakukannya penelitian ini dapat menjadi referensi bagi penelitian - penelitian yang sejenis.

## **1.6 Sistematika Penulisan**

Secara umum, penulisan skripsi ini terbagi dalam lima bab. Pembahasan yang terkandung dalam bab satu dengan bab lainnya saling berkaitan satu sama lain. Sehingga pada akhirnya akan membentuk suatu karya tulis yang runtut dan sistematis. Adapun sistematika penulisannya adalah sebagai berikut :

### **BAB I PENDAHULUAN**

Bab ini mencakup latar belakang masalah, rumusan masalah, ruang lingkup, tujuan penelitian, manfaat penelitian, dan sistematika penulisan dari penelitian yang dilakukan.

### **BAB II TINJAUAN PUSTAKA DAN DASAR TEORI**

Bab ini menjelaskan tinjauan pustaka untuk membedakan penelitian ini dari penelitian yang sebelumnya dan dasar teori yang dibutuhkan dalam proses analisis pembuatan sistem.

### **BAB III METODE PENELITIAN**

Bab ini berisi bahan/data dan peralatan yang dibutuhkan dalam proses analisis dan pembuatan sistem, prosedur dan pengumpulan data penelitian, serta analisis dan perancangan sistem yang akan dibuat.

### **BAB IV IMPLEMENTASI DAN PEMBAHASAN**

Bab ini berisi implementasi dan uji coba sistem yang membahas implementasi dari metode yang digunakan, kelebihan dan kekurangan yang diperoleh, serta pembahasan kajian/bahasan tentang hasil pengujian yang diperoleh dari penelitian.

### **BAB V PENUTUP**

Bab ini berisi kesimpulan yang berisi jawaban dari pertanyaan penelitian yang dinyatakan dalam perumusan masalah, dan saran yang perlu dilanjutkan atau direalisasikan.