

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Era digital saat ini membuat data memegang peranan penting dalam sebuah bisnis, industri, atau bahkan individu. Data memegang peranan penting karena merupakan kunci pengambilan keputusan baik dalam bisnis, industri, maupun individu. Saat ini, cara efektif dan efisien yang dapat digunakan untuk menyimpan dan mengelola data berukuran besar adalah dengan menggunakan *database*. *Database* adalah kumpulan informasi yang diorganisasikan dan disimpan di dalam komputer agar dapat dijalankan dengan program komputer untuk mendapatkan informasi [1]. *Database* dibuat dengan tujuan untuk mempermudah proses pencarian dan proses mendapatkan data. *Database* biasanya dapat dikontrol dengan *database management system* (DBMS) [2]. DBMS adalah *software* yang menyediakan fasilitas untuk mengakses data yang disimpan pada *database*.

RDBMS, atau *relational database management system* adalah salah satu jenis DBMS yang mengatur data ke dalam baris dan kolom, yang secara kolektif membentuk tabel dengan data tabel yang saling terkait [3]. RDBMS sendiri memiliki banyak keunggulan yang membuat *database* jenis ini banyak digunakan dalam pengerjaan *project* ataupun dalam perusahaan. Pertama, penggunaan RDBMS dapat menghindari duplikasi dengan menyimpan informasi terkait di tabel berbeda, menggunakan SQL. Selain itu, data dapat diakses dengan mudah dengan adanya *graphical user interface* (GUI) dan dapat menggunakan *Structured Query Language* (SQL) sebagai bahasa untuk berinteraksi, sehingga memungkinkan pertukaran data dalam RDBMS yang berbeda [4]. Popularitas dari *database* umumnya dapat dilihat melalui DB Engines *website* yang menyediakan *ranking database* terpopuler setiap tahunnya.

423 systems in ranking, January 2025

Rank			DBMS	Database Model	Score		
Jan 2025	Dec 2024	Jan 2024			Jan 2025	Dec 2024	Jan 2024
1.	1.	1.	Oracle	Relational, Multi-model	1258.76	-5.03	+11.27
2.	2.	2.	MySQL	Relational, Multi-model	998.15	-5.61	-125.31
3.	3.	3.	Microsoft SQL Server	Relational, Multi-model	798.55	-7.14	-78.05
4.	4.	4.	PostgreSQL	Relational, Multi-model	663.41	-2.97	+14.45
5.	5.	5.	MongoDB	Document, Multi-model	402.50	+2.12	-14.98
6.	7.	9.	Snowflake	Relational	153.90	+6.54	+27.98
7.	6.	6.	Redis	Key-value, Multi-model	153.36	+3.08	-6.03
8.	8.	7.	Elasticsearch	Multi-model	134.92	+2.60	-1.15
9.	9.	8.	IBM Db2	Relational, Multi-model	122.97	+0.19	-9.43
10.	10.	11.	SQLite	Relational	106.69	+4.97	-8.51

Gambar 1. 10 DBMS Terpopuler di Januari 2025 [5]

Gambar 1.1 merupakan gambar peringkat *database management system* (DBMS) yang paling populer pada bulan Januari 2025 pada *website* DB Engines Ranking. Dari sepuluh nama DBMS yang ditampilkan, tujuh DBMS termasuk ke dalam model RDBMS, sisanya terdapat satu model *document*, satu model *key-value*, dan satu tipe *multi-model*. Hal ini menunjukkan bahwa tipe RDBMS masih menjadi salah satu tipe *database* yang populer sampai tahun 2025. Tiga diantaranya adalah *database* MySQL, Microsoft SQL Server, dan PostgreSQL. DB Engines Ranking menempatkan *database* MySQL di peringkat kedua dengan skor 998.15, *database* Microsoft SQL Server di peringkat ketiga dengan skor 798.55, sedangkan *database* PostgreSQL berada di peringkat keempat dengan skor 663.41 [5].

Secara garis besar, MySQL, Microsoft SQL Server, dan PostgreSQL adalah *database* yang banyak digunakan oleh perusahaan dan didukung oleh komunitas pengguna yang besar. Walaupun ketiga *database* ini termasuk dalam kategori RDBMS, namun tetap memiliki beberapa perbedaan. Sebagai contoh, MySQL adalah RDBMS *open-source* yang dimiliki oleh Oracle, sedangkan PostgreSQL adalah RDBMS *open-source* yang dimiliki oleh PostgreSQL Global Development Group [6]. Selain itu, *database* PostgreSQL memiliki keunggulan seperti menyediakan beberapa opsi *indexing* untuk meningkatkan kemampuan menjalankan *query* yang lebih kompleks, sedangkan *database* MySQL memiliki keunggulan pada performa *read-heavy environments* dan banyak digunakan pada Linux, Apache, MySQL, PHP (LAMP) *stacks*, sehingga banyak dijadikan pilihan untuk Developer dalam mengembangkan sistem [7]. Microsoft SQL Server adalah *database* yang dikembangkan oleh perusahaan Microsoft dan memiliki banyak

kelebihan juga, seperti memiliki fleksibilitas untuk digunakan di berbagai penggunaan, mulai dari penggunaan *data warehouse*, *business intelligence*, penyimpanan data, dan pengembangan *website*, serta mudah untuk terintegrasi dengan produk Microsoft lainnya, seperti Microsoft Excel dan Power BI.

PT. GKT adalah salah satu perusahaan distribusi produk-produk yang berkaitan dengan listrik, seperti lampu, kipas angin, kabel, steker, dan lain-lain. Sampai saat ini, sudah ada lebih dari ratusan produk listrik yang dijual ke pelanggan setiap harinya, sehingga membutuhkan adanya *database* agar data dapat tersimpan dengan aman untuk kebutuhan jangka panjang. Data yang akan digunakan adalah data dari PT. GKT karena memiliki jumlah transaksi yang banyak, variasi produk yang beragam, dan proses bisnis secara keseluruhan yang tidak terlalu rumit, sehingga cocok untuk digunakan dalam kebutuhan penelitian, namun untuk penelitian ini, data yang digunakan nantinya akan digunakan yang berkaitan dengan transaksi penjualan saja karena jumlah transaksi penjualan yang dilakukan per-harinya dengan *customer* lebih banyak dibandingkan dengan pembelian dengan *supplier*. Umumnya, pembelian barang ke *supplier* dilakukan dengan jumlah barang yang sangat banyak untuk melakukan *restock* produk pada bagian *warehouse* apabila stok produk sudah hampir habis, sehingga untuk data jumlah transaksi pembeliannya tidak sebanyak jumlah transaksi penjualan.

Setiap RDBMS perlu untuk memiliki *response time* yang baik, terutama ketika proses pencarian data untuk memastikan data yang dimunculkan bisa muncul dengan cepat dan pengguna juga tidak perlu menunggu hasil datanya terlalu lama. Oleh karena itu, perlu untuk melakukan optimasi kinerja pencarian data pada *database* untuk menghasilkan *response time* yang lebih baik. Ada beberapa cara yang dapat dilakukan untuk mengoptimalkan kinerja *database*, contohnya adalah dengan mengoptimalkan *query* dengan struktur yang efisien, mengimplementasikan *cache* untuk menyimpan *query* yang sering digunakan, mengoptimalkan penggunaan *memory* pada *database*, dan lain-lain [8].

Secara garis besar, ada beberapa cara yang dapat dilakukan untuk mengoptimalkan kinerja *database* dari sisi struktur penulisan *query* SQL. Umumnya, proses optimasi dilakukan dengan dua cara, yaitu dengan *heuristic optimizations* dan *cost-based optimizations*, namun dalam penerapannya, kedua

proses ini dapat digabungkan untuk mendapatkan performa yang lebih maksimal pada *database*. Beberapa cara yang dapat dilakukan untuk meningkatkan performa *query* SQL dari sisi penulisannya adalah seperti memberikan *index database*, memberikan *partition table*, dan lain-lain. Setiap optimasi yang dilakukan untuk struktur *query* SQL lainnya yang juga membutuhkan pertimbangan juga dalam penulisan agar hasil performa *database*-nya bisa lebih optimal [9].

Dalam melakukan pengujian *database*, ada beberapa *tools* juga yang dapat digunakan sesuai dengan kebutuhan. Salah satu *tools* yang umum digunakan adalah Apache JMeter. JMeter sendiri adalah aplikasi pengujian performa berbasis *desktop* dari Apache yang ditulis dengan bahasa pemrograman Java, khususnya untuk *load testing*. Salah satu alasan terkait penggunaan JMeter yang populer adalah karena ringan dan mudah untuk diunduh di *desktop*. Selain itu, adanya *graphical user interface* (GUI) juga menjadi alasan tambahan dalam penggunaan JMeter sebagai alat pengujian, walaupun JMeter sendiri tetap bisa dijalankan dengan berbagai metode lain, contohnya seperti dengan *meter.sh* di *bash editor* [60].

Penelitian yang dilakukan ini bertujuan untuk mengetahui perbandingan performa RDBMS, terutama antara MySQL, PostgreSQL, dan Microsoft SQL Server ketika menjalankan *query* SQL sebelum dan sesudah menerapkan optimasi struktur penulisan *query*. Kebutuhan data yang digunakan saat ini adalah sampel data saja dengan jumlah baris data di sekitar 100.000 lebih data transaksi penjualan. Selain itu, hasil dari penelitian ini juga dapat digunakan membandingkan performa antara ketiga RDBMS yang memiliki kinerja terbaik setelah melakukan optimasi dari sisi penulisan *query*. Hasil performa RDBMS yang terbaik dari pengujian sebelum dan sesudah dioptimalkan diharapkan dapat menjadi pertimbangan perusahaan dalam menentukan *database* untuk kebutuhan penyimpanan data perusahaan.

## 1.2 Rumusan Masalah

Beberapa rumusan masalah yang digunakan dalam penelitian adalah sebagai berikut:

1. Bagaimana hasil *database performance* dari MySQL, PostgreSQL, dan Microsoft SQL Server sebelum menggunakan optimasi *query* dan sesudah menggunakan optimasi *query*?
2. Bagaimana cara membuat penulisan struktur *query* yang efektif sebagai salah satu cara untuk meningkatkan performa kecepatan pencarian data di *database*?

### 1.3 Batasan Masalah

Beberapa batasan masalah yang digunakan dalam penelitian adalah sebagai berikut:

1. Tipe *relational database management system* (RDBMS) yang akan digunakan di dalam penelitian hanya MySQL, PostgreSQL, dan Microsoft SQL Server saja.
2. Pengujian *query* yang dilakukan akan menggunakan SELECT *query* dengan total *query* yang diujikan adalah 10 *query*.
3. Optimasi pada *database* yang akan digunakan di dalam penelitian lebih difokuskan pada struktur penulisan *query* SQL.
4. Metrik performa *database* yang akan diuji dan dicek di dalam penelitian akan lebih difokuskan pada bagian *response time*, *throughput*, dan *error rate* di Apache JMeter.
5. Data yang digunakan yaitu data transaksi penjualan perusahaan karena memiliki baris data yang lebih banyak sehingga cocok digunakan untuk penelitian.
6. *Tools* yang digunakan untuk pengujian performa *database* adalah menggunakan Apache JMeter.

### 1.4 Tujuan dan Manfaat Penelitian

#### 1.4.1 Tujuan Penelitian

Terdapat dua tujuan penelitian yang ingin dicapai dalam penelitian, yaitu:

1. Melakukan analisis perbandingan performa RDBMS antara MySQL, PostgreSQL, dan Microsoft SQL Server sebelum dan sesudah

mengoptimalkan struktur penulisan *query database* untuk memberikan rekomendasi RDBMS dengan performa terbaik ke perusahaan.

2. Melakukan analisis terkait penulisan struktur *query database* yang efektif sebagai salah satu cara untuk meningkatkan performa kecepatan pencarian data di *database*.

#### 1.4.2 Manfaat Penelitian

Terdapat tiga manfaat dari hasil penelitian ini. Ketiga manfaat penelitian ini adalah sebagai berikut:

1. Memberikan pemahaman terkait penerapan struktur *query* yang efektif sebagai salah satu cara untuk mengoptimalkan kinerja *database*.
2. Memberikan pemahaman terkait aspek-aspek yang penting dalam pengecekan performa *database*.
3. Memahami perbedaan performa pada RDBMS MySQL, PostgreSQL, dan Microsoft SQL Server, terutama sebelum dan sesudah menggunakan penulisan struktur *query* yang lebih efektif.
4. Memberikan rekomendasi terkait penggunaan RDBMS yang memiliki performa terbaik ke PT. GKT sebagai perusahaan objek penelitian.

#### 1.5 Sistematika Penulisan

Sistematika penulisan adalah bagian yang diisi dengan penjelasan singkat mengenai gambaran isi penulisan laporan penelitian secara keseluruhan, mulai dari bagian pendahuluan hingga bagian terakhir, yaitu kesimpulan dan saran. Berikut ini adalah gambaran isi setiap bab pada penelitian:

##### 1. Bab 1

Bagian ini akan membahas terkait latar belakang penelitian ini dilakukan, rumusan masalah, batasan masalah, dan tujuan dilakukannya penelitian pengujian performa RDBMS.

##### 2. Bab 2

Bagian ini nantinya akan membahas kajian literatur dan juga penelitian terdahulu terkait *relational database* dan pengujian performa *database*. Selain itu, bab ini juga akan membahas terkait *database* yang digunakan

dalam penelitian dan *tools* yang akan digunakan untuk melakukan pengujian *database*.

### 3. Bab 3

Bagian ini nantinya akan membahas terkait langkah-langkah penelitian yang akan digunakan untuk melakukan penelitian dengan DSDLC.

### 4. Bab 4

Bagian ini nantinya akan membahas hasil penelitian pengujian *database management system* (DBMS) pada MySQL, PostgreSQL, dan Microsoft SQL Server. Setelah itu, hasil *query* dari ketiga *database* akan dibandingkan hasil performanya dengan *tools* pengujian yang sudah dipilih untuk penelitian.

### 5. Bab 5

Bagian ini nantinya akan membahas kesimpulan penelitian dari hasil penelitian. Kesimpulan yang dicantumkan disesuaikan dengan tujuan penelitian pada Bab 1. Selain itu, ada juga tambahan saran untuk penelitian-penelitian yang serupa.

