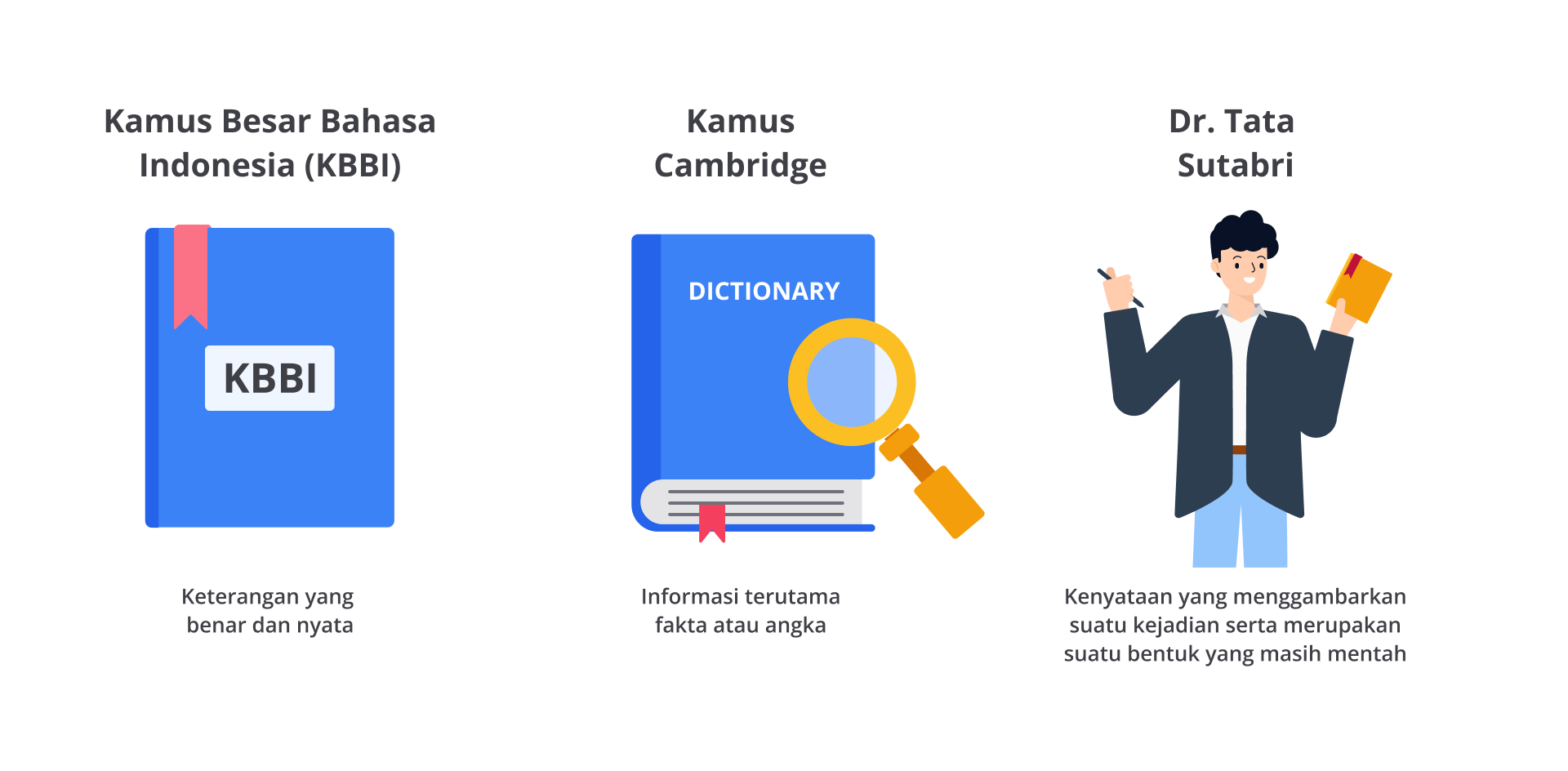
**Mengenal Data**

**Mengenali Data Secara Umum**

Menurut Kamus Besar Bahasa Indonesia, data merupakan keterangan yang benar dan nyata; keterangan atau bahan nyata yang dapat dijadikan dasar kajian; informasi dalam bentuk yang dapat diproses oleh komputer, seperti representasi digital dari teks, angka, gambar grafis, atau suara.

Menurut Kamus Cambridge, data adalah informasi, terutama fakta atau angka, dikumpulkan untuk diperiksa dan dipertimbangkan, serta digunakan untuk membantu pengambilan keputusan atau informasi dalam bentuk elektronik yang dapat disimpan dan digunakan oleh komputer.

Terakhir, menurut ahli, yaitu Tata Sutabri, dalam buku *Konsep Sistem Informasi*, data adalah kenyataan untuk menggambarkan suatu kejadian serta merupakan suatu bentuk yang masih mentah dan belum dapat bercerita banyak sehingga perlu diolah lebih lanjut melalui suatu model untuk menghasilkan informasi.



Dari ketiga pengertian di atas, dapat disimpulkan tiga kata yang mewakili arti sebuah data, yaitu Informasi yang benar, fakta, dan nyata. Dengan kata lain, data merupakan kenyataan apa adanya (*raw facts*).

**Pentingnya Data**

Anda adalah solo traveler tanpa pengalaman sebelumnya. Apakah Anda tersadar bahwa berbagai informasi yang telah Anda kumpulkan sebelumnya hingga hal-hal yang Anda lihat di tempat wisata tersebut semuanya adalah data?

Ketika mencari tahu terkait negara tujuan yang akan dikunjungi, Anda akan mendapatkan data terkait tempat tersebut. Petunjuk jalan yang Anda lihat di sepanjang jalan juga merupakan representasi dari data, lalu turis-turis yang Anda temui pun merupakan data yang bisa dikumpulkan untuk informasi wisatawan, dan juga makanan yang disantap bisa menjadi data yang dapat diolah sebagai informasi untuk orang-orang yang akan mengunjungi tempat wisata tersebut. Namun, ….

Terbayangkaholeh Anda apabila tidak ada hal yang dapat diketahui sebelumnya sebelum menjalani *solo traveling*? Tidak ada petunjuk jalan menuju destinasi wisata? Tidak ada informasi terkait makanan? Hal-hal yang Anda ketahui tersebut merupakan salah satu peran penting dari data. Bisa dikatakan, Anda akan buta terhadap hal yang akan dihadapi. Apakah terbayang apabila tidak ada data sedikit pun di dunia ini? Mungkin, Anda tidak tahu tahun lahir Anda, berat badan Anda sewaktu bayi, atau bahkan Anda tidak memiliki identitas.

Maka dari itu, dapat terlihat, bukan? Ternyata, data sangat dibutuhkan dan sangat amat penting ketika kita sedang menjalani aktivitas yang relevan.

**Metode Pengumpulan Data**

Sudah tahukah Anda?  Ada dua metode pengumpulan data, yaitu untuk data kualitatif dan data kuantitatif. Apakah Anda masih bingung bagaimana cara membedakan data kualitatif dan data kuantitatif? Yuk, kita simak materi berikutnya.

**Data Kualitatif**

Menurut Sugiyono (2014) data kualitatif adalah data yang berbentuk kata, skema, dan gambar. Data kualitatif didapatkan dari metode penelitian yang berlandaskan filsafat. Teknik pengumpulan data dan analisis bersifat kualitatif atau lebih menekankan pada makna.

Mari kita menganalogikan bahwa Anda akan pergi ke Gedung Sate di kota Bandung menggunakan kereta seperti yang terlihat pada gambar di atas. Namun, Anda ingin mencari testimoni dari teman-teman terlebih dahulu. Kita analogikan nama teman-teman Anda pada gambar di bawah ini beserta jawabannya.



Dari hasil di atas, pasti Anda akan berpikir terlebih dahulu sebelum memutuskan menggunakan kereta tersebut. Apabila mempertimbangkan *experience*yang akan didapatkan, tentu Anda tidak akan berpikir dua kali untuk menggunakan jenis tersebut. Namun, apabila memikirkan biaya yang akan dihabiskan, bisa jadi Anda lebih memilih naik kereta ekonomi dibandingkan dengan kereta panoramic tersebut.

Nah, data yang diperoleh dari teman-teman Anda tersebut termasuk data kualitatif karena bergantung pada kualitas objek. Maka dari itu, data kualitatif bisa bersifat subjektif karena hasilnya dapat diartikan dengan persepsi yang berbeda apabila dibaca atau dilihat orang lain. Pada contoh kasus di atas, Budi mengatakan sangat *worth it*tetapi Ica mengatakan hal yang bertolak belakang.

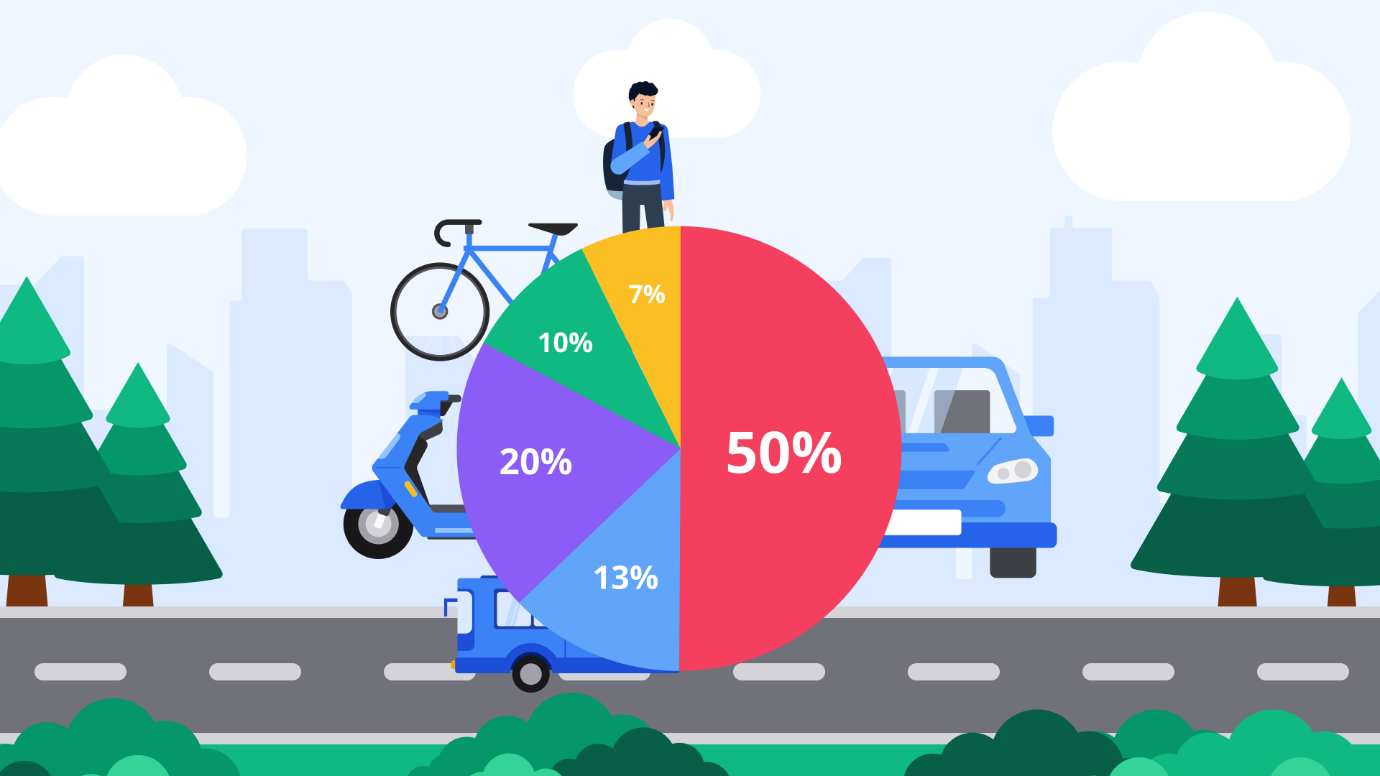
Dari kasus di atas, maka dapat disimpulkan bahwa kualitatif yaitu data yang bersifat verbal, dapat dengan mudah dicerna oleh manusia, sehingga lebih mudah dalam mengidentifikasi dan menghasilkan pemecahan masalah atau keputusan yang tepat.

Jadi, kalau Anda memiliki pilihan lebih memilih pergi ke Bandung menggunakan kereta panoramic atau beli kebutuhan lain? *Hehehe.*

**Data Kuantitatif**

Data kuantitatif merupakan data yang dinyatakan dalam bentuk angka [4]. Data ini bisa didapatkan dari hasil pengukuran atau observasi yang memerlukan instrumen jumlah data banyak dan tidak terlalu detailterhadap hasil pengukuran tersebut. Mari kita analogikan data kuantitatif.

Di bawah ini merupakan salah satu contoh data kuantitatif, yaitu berupa persentase yang telah dirangkai menggunakan *pie chart*,merepresentasikan pengguna transportasi umum di Indonesia.



Data di atas bisa didapatkan dari penelitian dengan membagikan kuesioner kepada para responden yang telah ditentukan populasi dan sampelnya terlebih dahulu.

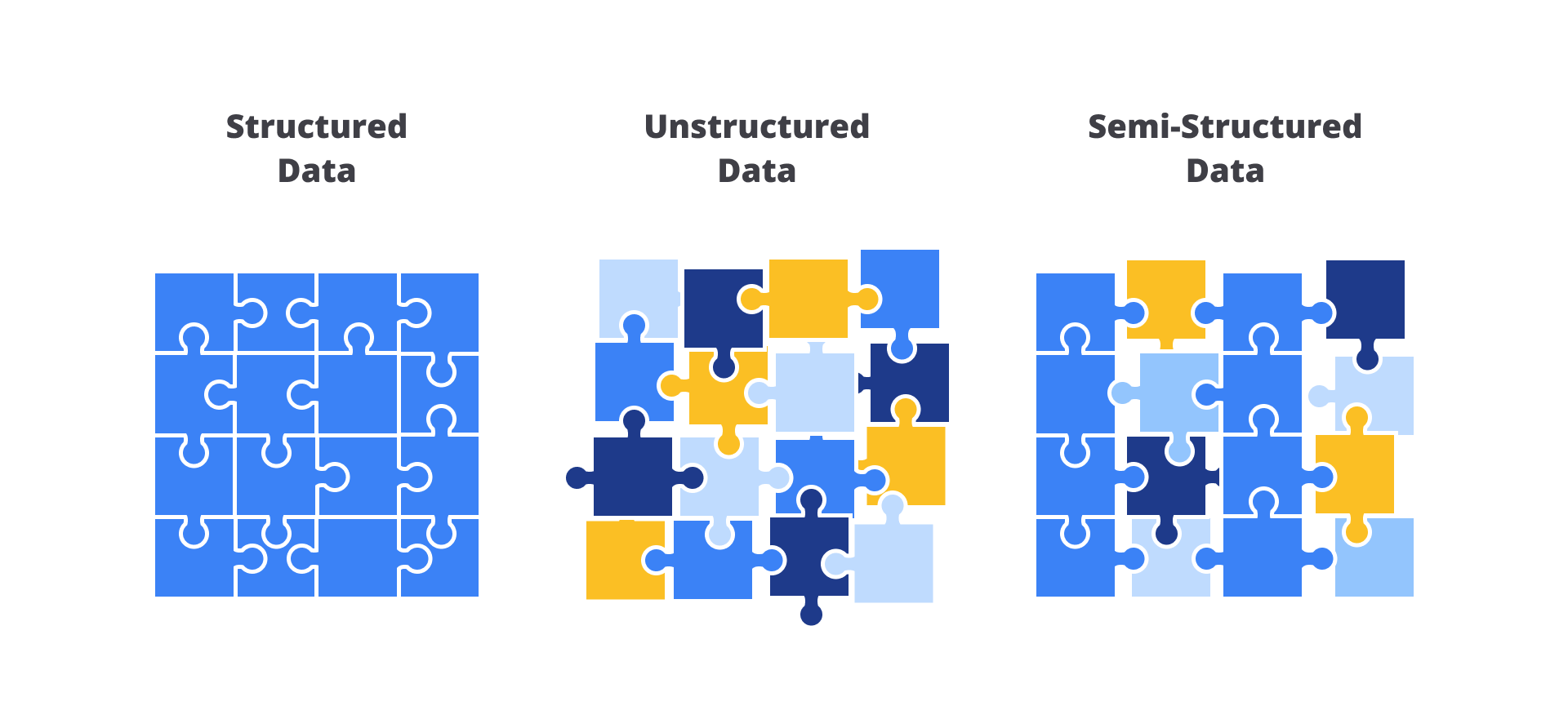
Nah, kumpulan data tersebut merupakan data kuantitatif. Data yang dikumpulkan bersifat objektif dan mutlak (ditunjukkan dengan angka).

**Jenis-Jenis Data**



Jenis-jenis data yang akan Anda pelajari di antaranya adalah data terstruktur, data tidak terstruktur, data semi terstruktur, data numerik, dan data kategorikal. Sangat menarik, bukan? Yuk, kita langsung ke materi berikutnya.

**Data Terstruktur, Data Tidak Terstruktur, Data Semi Terstruktur**



**Data Terstruktur**



Terlihat bahwa data tersebut tersusun rapi sesuai dengan struktur tabelnya. Jadi, orang dapat langsung memahami maksud yang ingin disampaikan dari tabel tersebut. Contohnya, terlihat jelas bahwa total penumpang yang memilih tujuan wisata ke Dago Pakar sebanyak 300 orang.

Jadi, dapat disimpulkan bahwa data terstruktur adalah salah satu jenis data yang disusun dengan rapi dan diatur sedemikian rupa sehingga memiliki format atau bentuk yang tetap.

**Data Tidak Terstruktur**

Seperti namanya, data tidak terstruktur, ia tidak memiliki struktur yang beraturan. Mari simak gambar di bawah ini.



Apakah Anda menemukan perbedaan antara data terstruktur dan tidak terstruktur? Sudah jelas pasti jawabannya sangat berbeda, ya.

Gambar di atas merupakan *flyer package sleeper bus*. Terlihat dari *flyer*tersebut terdapat beberapa format data, seperti gambar dan teks. Dari data tidak terstruktur, kita dapat mempelajari bahwa sebuah data tidak harus selalu berbentuk tabel dan angka, tetapi jauh lebih luas dari itu.

**Data Semi Terstruktur**

Data semi terstruktur adalah data dengan bentuk yang tidak dikenal. Pada dasarnya, jenis data ini memiliki atribut terstruktur yang membuatnya tetap dapat dianalisis. Namun, pada saat yang bersamaan, data semi terstruktur ini tidak mengikuti suatu susunan tertentu seperti data tidak terstruktur.

Contoh data semi terstruktur di antaranya adalah email, halaman web, file yang ada dalam sebuah *zipped*folder, dan sebagainya.

Nah, selain data terstruktur, tidak terstruktur, dan semi-terstruktur, ada pula jenis-jenis data lain yang dapat Anda ketahui, yaitu numerik dan kategorikal. Yuk,simak penjelasannya di bawah ini.

**Data Numerik dan Data Kategorikal**

Selain terstruktur dan tidak terstruktur, jenis-jenis data yang lain adalah numerik dan kategorikal. Mari kita simak perbedaannya di bawah ini.

**Data Numerik**

Data numerik adalah data berwujud angka yang bisa didapat dari sebuah pengukuran. Misal, ukuran tinggi badan, berat badan, dan tekanan darah kita. Selain itu, data ini juga bisa diperoleh dari sebuah penghitungan, misalnya jumlah orang yang hadir di pesta pernikahan. Nah, data numerik ini ternyata masih bisa dibagi menjadi beberapa jenis lagi, di antaranya data diskrit dan data kontinu.

**Data Kontinu**

Data kontinu dapat direpresentasikan dalam berbagai nilai numerik, seperti bilangan desimal, bulat, dan lain-lain. Contohnya, setelah datangnya pandemi, saat Anda akan memasuki kapal feri menuju Banda Neira, Anda diminta untuk cek suhu terlebih dahulu. Setelah dilihat suhu Anda menunjukan 36.2 derajat Celcius. Lalu, data tersebut dicocokkan dengan persyaratan masuk ke dalam feri untuk dikategorikan.

**Data Diskrit**

Data diskrit merupakan data numerik yang hanya bisa direpresentasikan dengan bilangan bulat dan tidak dapat dibagi ke dalam unit yang lebih kecil. Mari kita ambil contoh. Setelah Anda diperiksa suhu ketika memasuki kapal feri, Anda melihat ternyata jumlah penumpang kapal menuju Banda Neira sebanyak 125 orang.

**Data Kategorikal**

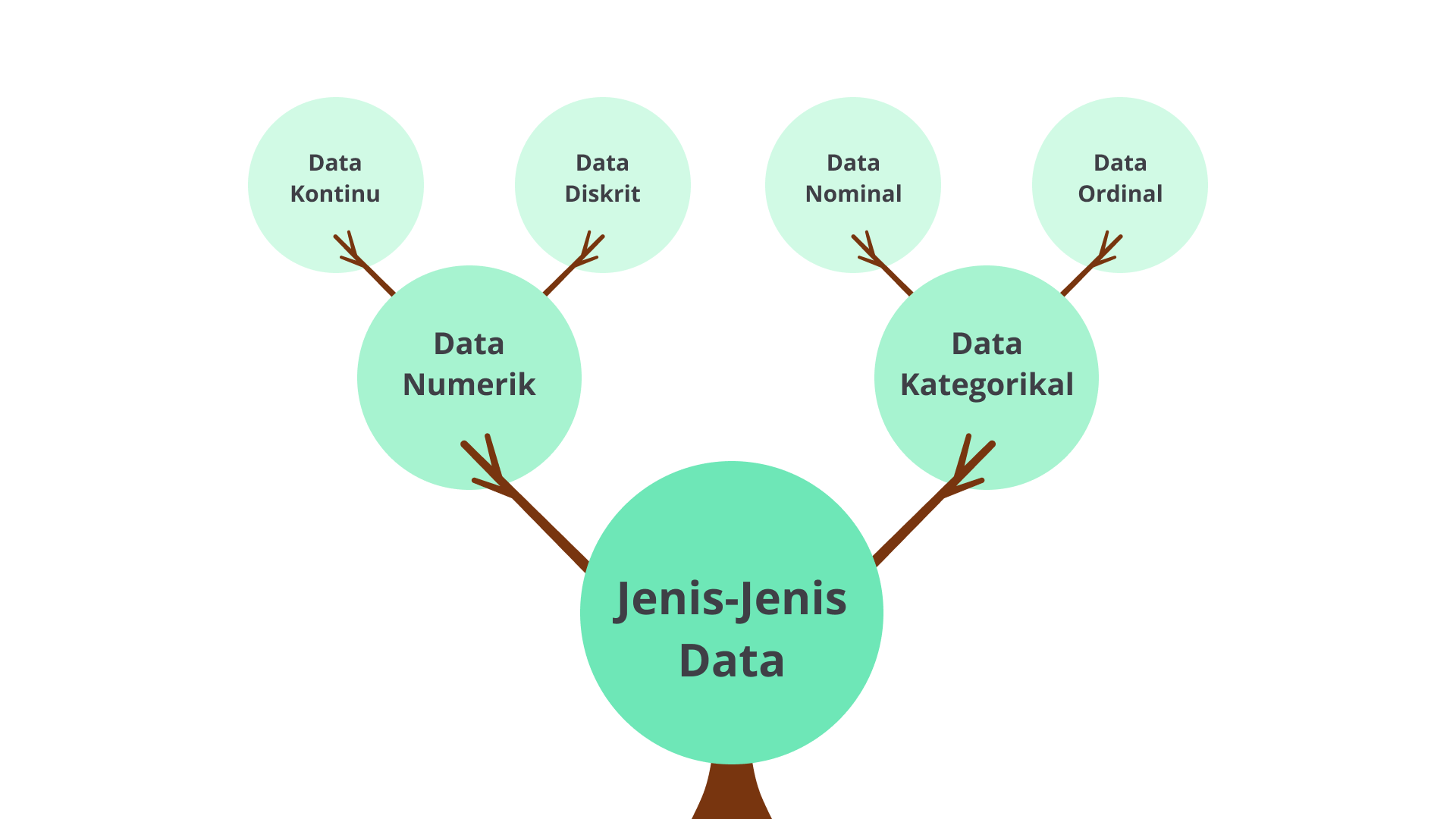
Data kategorikal merupakan data yang dapat dikelompokkan dan terbagi berdasarkan karakteristik atau ciri khasnya masing-masing. Dari data kategorikal, ada dua pembagian, yaitu nominal dan ordinal.

**Data Nominal**

Data nominal adalah jenis pengelompokan data yang tidak memiliki keterkaitan dengan data lainnya dan tidak memiliki arti khusus. Jadi, data ini dapat dibedakan tanpa harus mengurutkan atau dibandingkan dengan data lainnya. Pada dasarnya kendaraan bermotor dapat diklasifikasikan menjadi transportasi publik, tetapi antar kendaraan tersebut tidak terikat satu sama lain dan tidak memiliki arti khusus.

**Data Ordinal**

Berlawanan dari data nominal, data ordinal adalah jenis pengelompokkan data yang memiliki urutan, atau harus disusun secara berurutan dengan mekanisme peringkat. Misalnya, kita mengambil voting destinasi laut terbaik di Indonesia, seperti gambar di bawah ini. Manakah menurut Anda yang lebih bagus? Apakah Raja Ampat, Papua? Atau Banda Neira, Kepulauan Banda? Berdasarkan gambar yang Anda lihat bahwa terdapat *rating* atau peringkat pada kedua destinasi liburan. Bintang 1 mengartikan bahwa destinasi tersebut kurang peminat, sedangkan bintang 5 mengartikan bahwa banyak wisatawan yang puas akan destinasi wisata tersebut.



**Tipe-Tipe Data**

Tipe-tipe data merupakan hal penting yang perlu Anda pelajari sebelum memulai pemrograman karena hal ini akan menjadi rutinitas yang akan Anda temui setiap mempelajarinya. Tentunya, banyak sekali tipe-tipe data dalam bahasa pemrograman seperti *Integer*, *Character, Boolean*, *Array*, *String*, dan *Float*.

**Tipe Data yang Sering Digunakan**

Tabel di bawah ini merupakan beberapa tipe data yang mungkin akan Anda sering temukan ketika mempelajari bahasa pemrograman Structured Query Language.

* Integer (INT)

*Integer*(int.) merupakan tipe databerbentuk bilangan bulat atau numerik yang umumnya digunakan untuk menyimpan angka tanpa komponen pecahan. Tipe data ini mencakup semua bilangan bulat baik yang positif maupun negatif dengan *range*tertentu.

* Character (CHAR)

Karakter merupakan tipe data yang mengandung tabel ASCII, seperti huruf alfabet, simbol, tanda baca, serta emoji. Selain alfabet, Char dapat berupa huruf dari bahasa lain, seperti arabic, china, jepang, dll.

* Boolean (Bool)

*Boolean*adalah jenis tipe data untuk mewakili nilai yang benar dan salah dalam data.

* Array

*Array* adalah tipe data untuk menyimpan sejumlah elemen dari urutan tertentu, dan biasanya berasal dari tipe yang sama.

* String (STR)

*String*merupakan tipe data yang paling umum digunakan untuk menyimpan teks. Selain itu, *string*juga dapat menyertakan angka dan simbol, tetapi ia akan selalu diperlakukan sebagai teks.

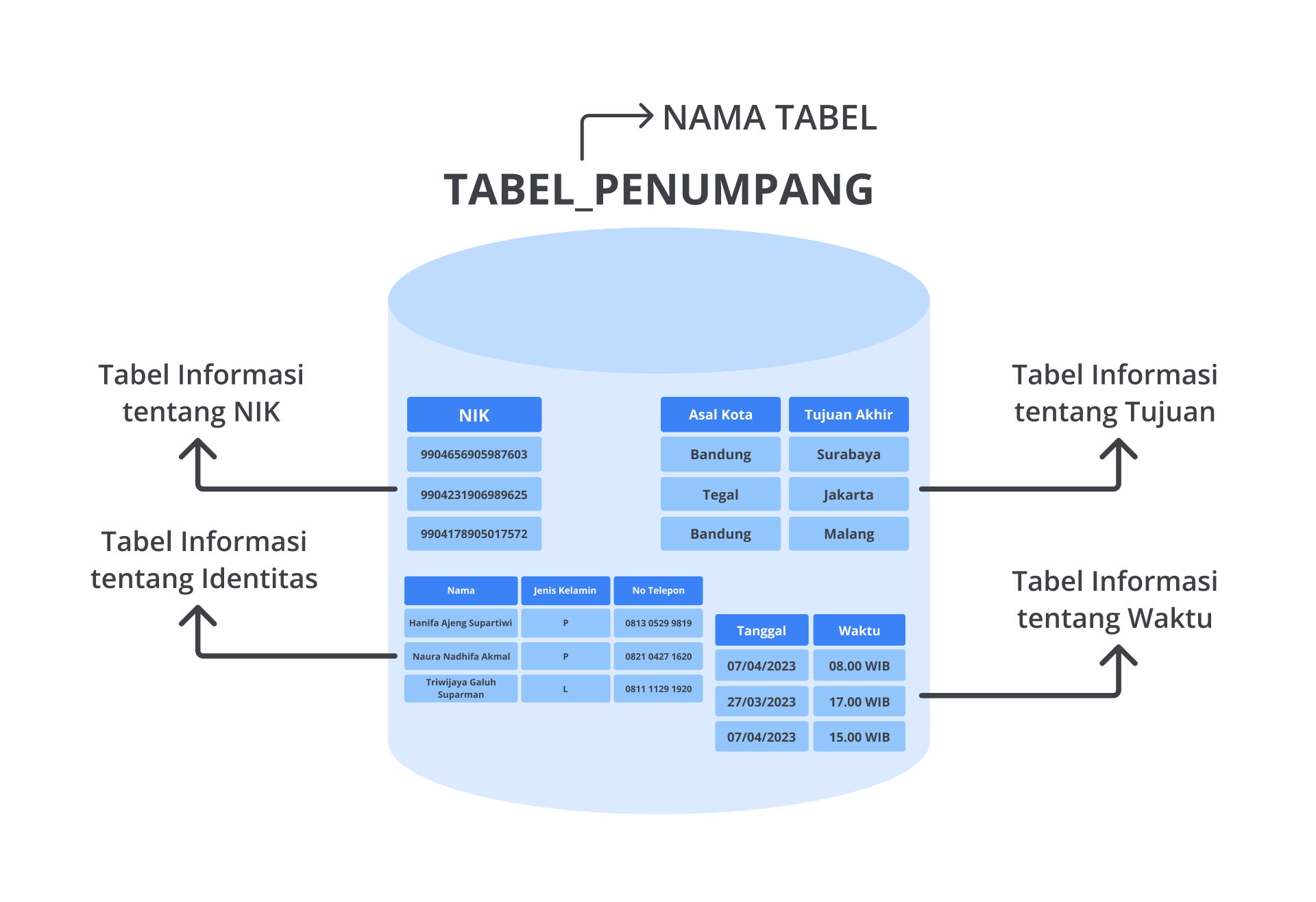
* Float

Float merupakan tipe data numerik untuk menyimpan angka yang mungkin memiliki komponen pecahan.

**Data, Basis Data, Dataset**

Telah dijelaskan di awal bagian kelas bahwa data didefinisikan sebagai informasi, fakta, dan nyata mengenai suatu objek atau kejadian yang bisa dicatat atau disimpan. Selain data dan basis data, ada hal yang perlu Anda tahu mengenai data, yaitu dataset. Dataset adalah kumpulan data yang disusun secara terstruktur. Biasanya, dataset dipresentasikan dalam bentuk tabel, alias kumpulan baris dan kolom. Simak penjelasannya pada tabel di bawah ini untuk mengetahui perbedaan yang jelas antara dataset dan basis data.

|  |  |
| --- | --- |
| **Dataset** | **Basis Data** |
| Dataset adalah kumpulan data yang diatur dalam format tertentu, seperti spreadsheet, CSV, atau basis data. Umumnya, digunakan untuk keperluan yang spesifik. | Basis data adalah kumpulan data yang diatur dan disimpan dengan cara yang memungkinkan akses dan pengambilan yang mudah. Serta dapat menyimpan berbagai macam tipe data, termasuk teks, nomor, gambar, dan tipe data lainnya. |
| Dapat menjadi bagian dari data yang diambil dari basis data yang lebih besar. Bisa juga, dataset diambil dari sumber lain yang mendukung formatnya. | Dapat memiliki banyak kumpulan data dan dapat digunakan untuk aplikasi yang berbeda. |
| Contoh tools: Google Spreadsheet, Excel, Situs Public Datasets, dll. | Contoh tools: BigQuery, MySQL Clients, DB Browser SQL, dll. |

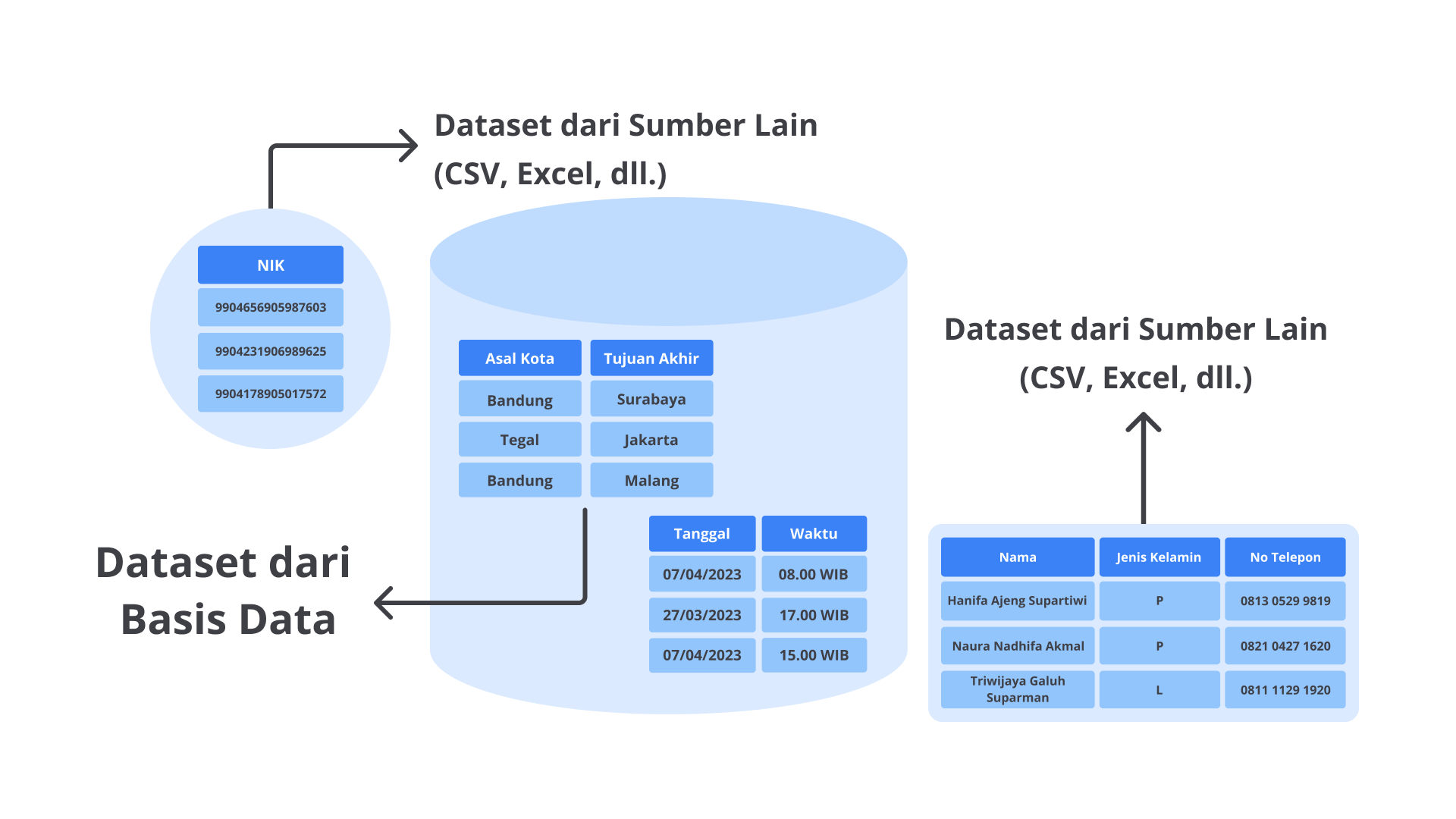


Gambar tabung di atas merupakan representasi dari sebuah basis data. Tabel yang terdapat di dalamnya adalah struktur dalam basis data yang diatur dengan bentuk kolom dan baris.

Ingat kategori yang terdapat pada gambar di atas? Setiap kategori menjadi kolom pada tabel tersebut. Sementara itu, baris berisi semua informasi tentang satu objek dalam tabel tersebut. Pada tabel yang telah dijelaskan dalam gambar, satu baris adalah tentang identitas seseorang, tujuan, dan waktu keberangkatan.

Selain hanya terdiri dari tabel, basis data juga terdiri dari data-data yang saling terhubung.

Namun, apabila tabel-tabel tersebut dibungkus dengan menggunakan format tertentu, ia disebut sebuah dataset. Dataset dapat Anda dapatkan baik dari database maupun dari sumber yang relevan dengan objek yang sedang dikerjakan. Perhatikan gambar di bawah ini agar Anda paham mengenai dataset.



Bagaimana menurut Anda? Apakah sampai sini paham perbedaan terkait data, basis data, dan dataset?

Mari simak analogi basis data di bawah ini agar Anda lebih paham.

Apabila dianalogikan, basis data adalah semacam hanggar, yaitu tempat penyimpanan pesawat dengan berbagai jenis, seperti helikopter, pesawat *airbus*, pesawat *boeing*, jet, dan lain-lain. Analogi tersebut serupa dengan definisi basis data yaitu data yang diatur dan disimpan dengan cara pengambilan yang mudah, serta menyimpan berbagai macam tipe data.

Adapun analogi lain adalah bahwa di dalam hanggar, ada berbagai macam pesawat dengan ragam jenis tertentu yang memiliki fungsinya masing-masing. Bisa juga digunakan untuk keperluan khusus, yaitu mengantarkan penumpang atau mungkin hanya sekadar untuk latihan terbang saja. Analogi tersebut selaras dengan definisi basis data bahwa ia memiliki banyak kumpulan data dan dapat digunakan untuk aplikasi yang berbeda-beda.

Lalu, analogi dataset sebagai berikut.



Mari kita analogikan bahwa dataset merupakan sebuah *showroom*dan format-format dalam basis data adalah jenis-jenis kendaraan yang terdapat dalam *showroom*, seperti mobil, motor, truk, dan lain-lain, serta mari kita analogikan basis data sebagai pabrik pembuatannya.

Ketika Anda mengunjungi *showroom,*Anda akan dihadapi oleh banyak jenisnya kendaraan, seperti *showroom*khusus mobil, motor, truk, dan lain-lain. Suatu hari Anda ingin membeli mobil tidak melalui *showroom,*tetapi langsung dari pabrik pembuatannya yang lebih besar. Serupa dengan definisi dataset yang dapat menjadi bagian dari data yang diambil dari basis data yang lebih besar.

Bagaimana menurut Anda? Apakah sampai sini sudah paham? Apabila masih belum paham, Anda bisa bertanya di Forum Diskusi, ya. *See you there!*

Di bawah ini merupakan beberapa platform yang bisa Anda gunakan sebagai referensi dalam mencari dataset.

1. **Kaggle**  
   Kaggle adalah sebuah komunitas yang menampung para pegiat *data science*yang ingin belajar lebih dalam. Saat ini, Kaggle telah memiliki lebih dari 1000 dataset beragam secara internasional, 170.000 post di forum. Berikut adalah kegiatan-kegiatan utama dan fasilitas yang bisa Anda dapatkan dari platform tersebut.
   * Data  
     Dataset-dataset ini dapat memudahkan Anda saat membuat model proyek yang ingin Anda kembangkan. Format file CSV adalah yang paling banyak ditemukan di Kaggle, tetapi ada juga datasetyang tersedia dalam format JSON, SQLite, *archive,*dan BigQuery.
   * Code  
     Ada banyak sekali jumlah kode tersedia di Kaggle. Kode-kode ini adalah hasil kontribusi dari pengguna yang tergabung di Kaggle. Tentunya, tersedianya sumber ini bisa menjadi bahan belajar *machine learning*yang mudah dan baik.
   * Komunitas  
     Tentunya di sini, Anda dapat membangun *network* dan bekerja bersama.
2. **Open Data Jabar**  
   Open Data Jabar adalah portal transparansi data dan informasi publik bagi warga Jawa Barat. Platform tersebut dapat diakses oleh siapa saja yang ingin menjelajahi dan mendapatkan koleksi dataset terbaru dari perangkat daerah di Jawa Barat. Format file yang bisa Anda dapatkan di platform ini adalah dalam bentuk CSV dan Excel.
3. **Badan Pusat Statistik**  
   Badan Pusat Statistik adalah Lembaga Pemerintah Non Kementerian yang bertanggung jawab langsung kepada Presiden. Portal BPS menyediakan dataset beragam dengan lingkup nasional. Hanya saja, format yang disediakan oleh BPS hanya berbentuk Excel. Jadi, apabila Anda akan menggunakan basis data, kemungkinan besar Anda harus mengubahnya terlebih dahulu.

**Mengenal Basis Data Relasional**

**Basis Data Relasional**

Basis data relasional pada dasarnya adalah cara menyusun informasi dalam tabel, baris, dan kolom atau kumpulan item data dengan hubungan yang telah ditentukan sebelumnya. Perhatikan contoh basis data relasional di bawah ini.



Pada tabel di atas, ada beberapa item yang akan dibutuhkan untuk pembuatan aplikasi biro perjalanan. Tabel-tabel di atas terdiri dari beberapa kategori, seperti objek wisata, destinasi, wisatawan, dan sirkulasi atau kumpulan dari seluruh tabel yang saling berelasi.

Item-item tabel tersebut, pada dasarnya tersusun dari kolom dan baris seperti yang tertera di bawah ini.

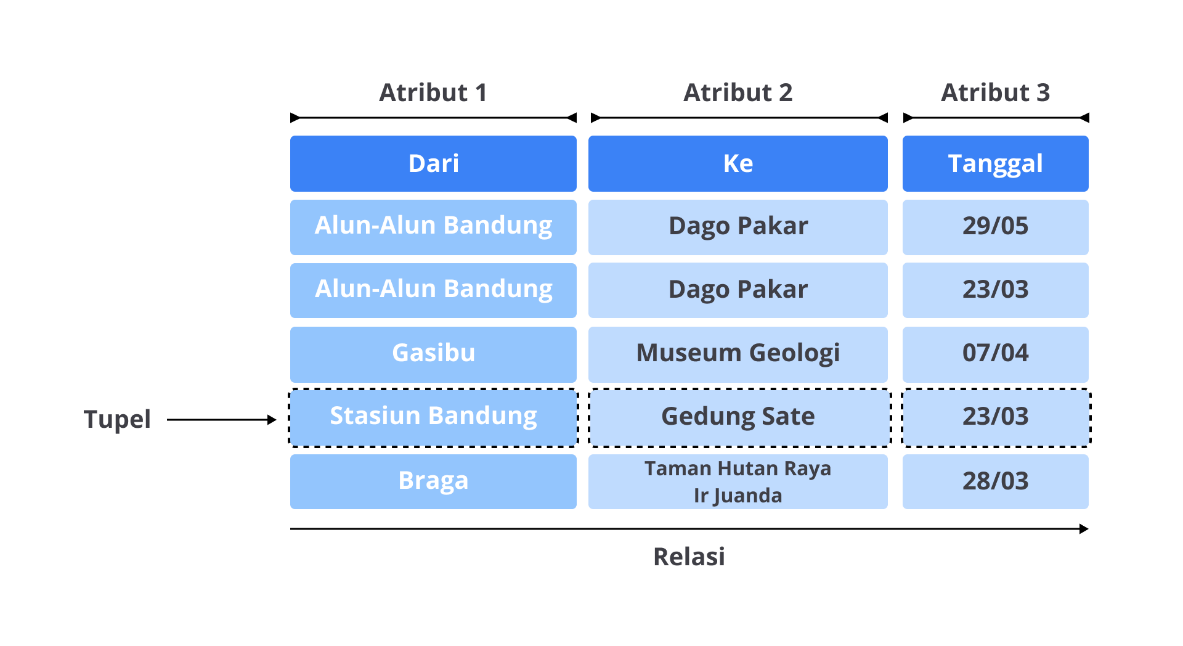


Seperti yang Anda lihat bahwa tabel tersebut terbentuk dari tiga kolom dan lima baris. Kolom adalah bagian dari data yang disimpan oleh tabel, sedangkan baris adalah satu set kolom yang menggambarkan atribut dari satu hal. Kolom dan baris bersama-sama membentuk tabel.

**Catatan:**

1. Baris pada tabel merepresentasikan kumpulan nilai terkait dari satu objek atau entitas.
2. Tiap baris pada tabel dapat ditandai dengan pengidentifikasi untuk yang disebut **Primary Key**, dan baris di antara beberapa tabel dapat dibuat saling terkait menggunakan **Foreign Key** (Akan dijelaskan pada kelas berikutnya, di kelas ini kita akan fokus pada penggunaan tabel tanpa relasi dengan tabel lain).
3. Data tersebut dapat diakses dengan berbagai cara tanpa menyusun ulang tabel basis data itu sendiri.

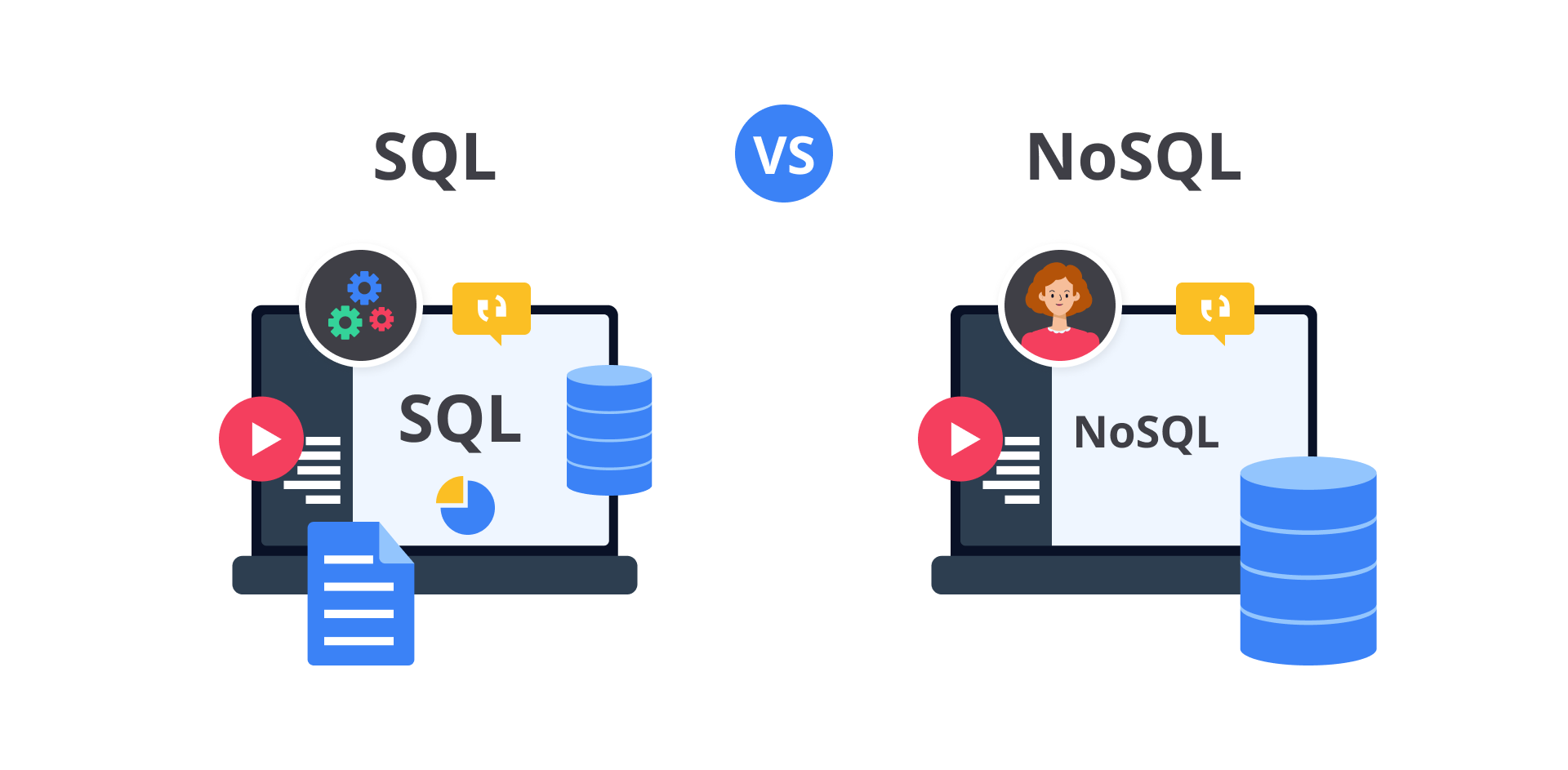
Di bawah ini merupakan istilah-istilah dalam basis data relasional. Perhatikan gambar berikut terlebih dahulu.



1. **Relasi (*Relation*)**  
   Relasi atau tabel terdiri dari beberapa kolom dan beberapa baris. Relasi menggambarkan beberapa atribut atau entitas-entitas di dalamnya. Contohnya, pada tabel destinasi di atas tertera bahwa tabel tersebut memiliki tiga kolom yang terdiri dari atribut 1, atribut 2, dan atribut 3 yang merepresentasikan nilai berbeda-beda.
2. **Derajat (*Degree*)**  
   Derajat merupakan jumlah atribut dalam sebuah relasi. Masih merujuk pada tabel di atas, berarti tabel tersebut memiliki derajat tiga karena terdiri dari tiga atribut.
3. **Tupel (*Tuple*)**  
   Baris pada sebuah relasi. Sebelumnya, Anda telah mempelajari perbedaan antara baris dan kolom. Setiap baris merujuk satu objek data. Perhatikan kembali pada tabel di atas, tabel tersebut memiliki lima objek data.
4. **Kardinalitas (*Cardinality*)**  
   Jumlah tupel dalam sebuah relasi. Tiap-tiap baris tentunya memiliki atribut yang sejenis. Merujuk pada tabel di atas, berarti tabel tersebut memiliki lima kardinalitas.

**Basis Data NoSQL**

NoSQL atau Not Only SQL adalah sebuah mekanisme penyimpanan dan pengambilan data, dalam pengambilan data tersebut tidak dibatasi dan memiliki konsistensi yang lebih baik dibandingkan dengan basis data relasional, serta tidak dibatasi tipe data dan variasi data.



Sebelum mengenal lebih jauh terkait perbedaan antara SQL dan NoSQL, di bawah ini merupakan hal yang perlu diketahui ketika Anda membutuhkan NoSQL.

* **Basis Data Relasional tidak restriktif.** Ketika basis data relasional sudah tidak bisa menangani data yang bersifat restriktif, yaitu data yang terdistribusi atau data yang bermacam-macam variasinya, maka solusinya adalah menggunakan NoSQL.
* **Ketika ACID sudah tidak dibutuhkan.**Dalam konteks pemrosesan transaksi, terdapat ACID yang mengacu pada empat sifat utama transaksi, yaitu

***Atomicity :*** Semua perubahan pada data dilakukan seolah-olah itu adalah satu operasi. Misalnya, dalam aplikasi yang mentransfer dana dari satu akun ke akun lainnya, properti atomocity atau atomisitas memastikan bahwa, jika debit berhasil dibuat dari satu akun, kredit yang sesuai dibuat ke akun lainnya.

***Consistency :*** Data berada dalam status konsisten saat transaksi dimulai dan diakhiri. Misalnya, dalam aplikasi yang mentransfer dana dari satu akun ke akun lainnya, properti konsistensi memastikan bahwa nilai dana di kedua akun sama di awal dan di akhir setiap transaksi.

***Isolation :*** Keadaan perantara dari suatu transaksi tidak terlihat oleh transaksi lain. Misalnya, dalam aplikasi yang mentransfer dana dari satu akun ke akun lainnya, properti isolasi memastikan bahwa transaksi lain melihat dana yang ditransfer di satu akun atau akun lainnya, tetapi tidak di keduanya, atau tidak keduanya.

***Durability :*** Setelah transaksi berhasil diselesaikan, perubahan pada data tetap ada dan tidak dibatalkan, bahkan jika terjadi kegagalan sistem. Misalnya, dalam aplikasi yang mentransfer dana dari satu akun ke akun lainnya, properti durability atau daya tahan memastikan bahwa perubahan yang dilakukan pada setiap akun tidak akan dibatalkan.

* Ketika ACID sudah tidak dibutuhkan, sistem yang tidak memenuhi kriteria ACID disebut dengan BASE, yaitu *Basically Available, Soft State,*dan *Eventual Consistent.*Sebagian besar NoSQL saat ini adalah BASE.
* **Adanya hambatan antara objek untuk berelasi.**Ketika antar objek tidak dapat berelasi karena banyak faktor dan hal tersebut menjadi hambatan, maka saatnya Anda menggunakan NoSQL.
* **RDBMS sudah tidak berdistribusi dan tidak scalable.**Artinya ketika basis data relasional sudah tidak bisa dikembangkan lagi karena tidak bisa dibuat menjadi sistem terdistribusi.
* **Dipakai ketika kita mengambil data yang sifatnya temporal (sementara).**Temporal data di antaranya adalah *shopping chart*dan *wish list.*Contoh *shopping chart*adalah ketika Anda berjualan, dan stoknya habis, maka data tersebut merupakan data temporal. Kemudian, ada *wish list* belanja pada sebuah marketplace, ketika *wish list*tersebut sudah tercapai maka data tersebut akan hilang dari *wish list.*

Lalu, apakah perbedaan yang menonjol antara SQL dan NoSQL?

|  |  |  |
| --- | --- | --- |
|  | SQL | NoSQL |
| Tipe Basis Data | SQL pada dasarnya berbasis Sistem Manajemen Relasional Basis Data atau *Relational Database Management System*(RDBMS), yaitu mereka menyimpan data dalam tabel dengan kolom yang telah ditentukan sebelumnya dan bergantung pada skema tabel tetap untuk menghubungkan catatan data satu sama lain. | Basis data NoSQL adalah basis data non-relasional (terdistribusi) jadi tidak memerlukan tabel. Maka dari itu, basis data non-relasional dapat menyimpan data tidak terstruktur menggunakan dokumen, nilai kunci, grafik, atau kolom. |
| Jenis Skala | Basis data SQL memiliki jenis skala vertikal. Ini berarti bahwa kebanyakan basis data ini hanya menggunakan satu server. Untuk menambahkan skala, Anda dapat menambahkan RAM atau CPU atau SSD yang mengakibatkan bertambahnya biaya untuk *upscale*SQL. Hal tersebut merupakan cara satu-satunya untuk dapat meningkatkan kapasitas server agar dapat menyimpan lebih banyak data.  Dengan catatan, banyak pengembangan tambahan yang membuat basis data SQL juga mampu lebih dari skala vertikal. | Basis data NoSQL dapat diskalakan secara horizontal. NoSQL dapat menggunakan banyak server dan berbagi data di seluruh server sehingga dapat menangani lebih banyak data. NoSQL menggunakan *hardware*berbiaya rendah. Ini membuatnya lebih mudah untuk *upscale*karena jumlah data bertambah dan basis data NoSQL bisa menjadi jauh lebih besar dibandingkan dengan basis data SQL yang hanya bergantung pada satu server. |
| Kapasitas Volume Data | Karena terbatasnya penyimpanan, basis data SQL akan mengalami kapasitas data yang penuh. Satu server hanya dapat menampung data sebelum mencapai kapasitas maksimumnya. Membeli server yang semakin besar, menjadikan SQL menjadi basis data yang tergolong mahal apabila ingin memiliki kapasitas yang besar. | Di sisi lain, basis data NoSQL dapat menangani volume data yang lebih tinggi karena dapat mendistribusikan data-data tersebut ke ribuan server yang terhubung. Untuk sebuah organisasi/perusahaan yang perlu menyimpan data tidak terstruktur dalam jumlah besar, basis data NoSQL adalah pilihan yang jelas. |

**Aljabar Relasional**

Sebenarnya basis data relasional memiliki hubungan dengan aljabar relasional yang berasal dari aljabar matematika, *lho.* Seperti *sigma, phi, rho,*dan masih banyak lagi. *First thing first,relational algebra*atau aljabar relasional merupakan kumpulan operasi terhadap relasi dan setiap operasi menggunakan satu atau lebih relasi untuk menghasilkan satu relasi yang baru.

Pada dasarnya, operasi aljabar relasional terdiri dari *selection, projection, cartesian, union, set-difference, dan rename.* Adapun operasi lain yang merupakan turunan aljabar relasional, yaitu set intersection, theta join, natural-join, outer join, dan division.



Operasi-operasi di atas terbagi kembali menjadi dua jenis, yaitu operasi *unary*dan operasi *binary.* Simak penjelasannya di bawah ini.

* Operasi *unary*terdiri dari *selection, projection.*Ia disebut operasi *unary*karena dapat digunakan hanya pada satu tabel.
* Operasi *binary* terdiri dari *union, intersection, set difference, cartesian product, join,*dan *division.*Ia disebut operasi *binary*karena memerlukan pasangan/kumpulan tabel.

Nah, di kelas dasar ini Anda hanya akan mempelajari terkait operasi *unary*yang hanya terdiri dari *selection*dan *projection.*Namun, jangan khawatir! Anda akan mempelajari operasi *binary*dengan operasi-operasi lainnya dalam pembahasan kelas selanjutnya.

**Operasi Selection**

*Selection*atau *select*() adalah operasi untuk menyeleksi tupel-tupel yang memenuhi suatu predikat. Kita dapat menggunakan operator perbandingan (<,>,>=,<=,=,#) pada predikat.

Pada dasarnya, operasi *selection*atau *select*ini berasal dari persamaan sederhana matematika yang diaplikasikan untuk kebutuhan query yang memiliki lambang *sigma*().

Beberapa predikat dapat dikombinasikan menjadi predikat majemuk (banyak) menggunakan penghubung AND (Λ) dan OR (V).

**Operasi Projection**

*Projection* atau *project*adalah operasi untuk memperoleh kolom-kolom tertentu. Operasi *project*termasuk dalam operasi *unary*yang mengirim relasi argumen dengan kolom-kolom tertentu. Sebab relasi himpunan, baris-baris duplikasi dihilangkan.

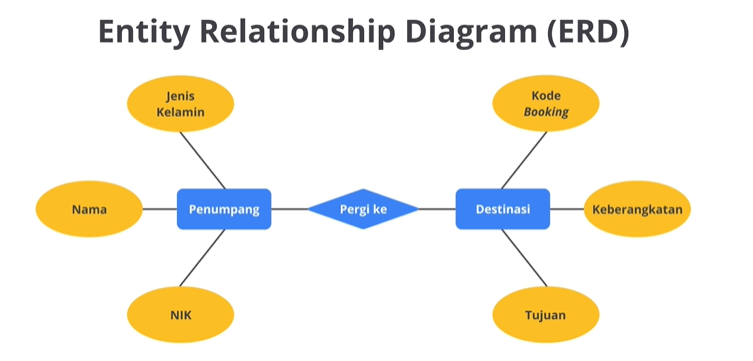
**Entity Relationship Diagram (ERD)**

*Entity Relationship Diagram* (ERD) merupakan salah satu teknik dalam merancang basis data. Dalam mempelajari ERD ada tiga hal yang perlu diketahui, yang pertama yaitu *entity. Entity*adalah terkait suatu objek/kejadian yang akan dicatat dan disimpan. Misalnya, dalam contoh yang termasuk dalam *entity* di antaranya penumpang, destinasi, masinis. Satu *entity*merepresentasikan satu objek, yang berarti satu penumpang. Lalu, kumpulan dari nama penumpang disebut *entity set* atau kumpulan objek-objek yang serupa.

Kedua, yaitu atribut atau properti, hal-hal yang mencirikan suatu objek. Contoh, apabila objeknya penumpang, atributnya bisa berupa nama, jenis kelamin, nomor induk keluarga.

Terakhir adalah relasi, yaitu hubungan antara *entity* dengan *entity* atau objek dengan objek, contohnya satu penumpang hubungannya dengan satu masinis, ia disebut *relationship set.*

Perlu digarisbawahi bahwa relasi di sini berbeda dengan relasi yang terdapat pada basis data. Agar lebih jelas, simak penjelasannya pada gambar di bawah ini agar Anda lebih paham dengan konsep ERD.



*Setiap entity*dilambangkan dengan bentuk persegi panjang. Pada diagram tersebut, *entity*satu adalah penumpang yang terdiri dari beberapa atribut yang dilambangkan oleh bentuk oval, yaitu nama, jenis kelamin, dan nomor induk keluarga.

Pada lambang kedua merupakan *entity* kedua yang merepresentasikan destinasi dengan atribut kode *booking*, keberangkatan, dan tujuan. *Entity* satu dan *entity* dua dihubungkan oleh relasi atau *relationship*yang digambarkan dengan bentuk belah ketupat. Bentuk belah ketupat adalah representasi dari relasi atau *relationship*antar kedua diagram tersebut.

Satu hal yang perlu ditekankan pada bagian relasi tersebut adalah “Apakah hubungan antara kedua diagram?” Jawabannya dapat Anda tulis pada diagram tersebut.

Nah, setelah menentukan hubungan antar diagram tersebut, hal lain yang harus diperhatikan adalah Anda perlu menentukan *primary key* dan *foreign key* yang akan dibahas pada modul berikutnya.