**Memilih data yang tepat**

Berikut adalah beberapa pertimbangan pengumpulan data yang perlu diingat untuk analisis Anda:

**Bagaimana data akan dikumpulkan**

Putuskan apakah Anda akan mengumpulkan data menggunakan sumber daya Anda sendiri atau menerima (dan mungkin membelinya) dari pihak lain. Data yang Anda kumpulkan sendiri disebut data pihak pertama.

**Sumber data**

Jika Anda tidak mengumpulkan data menggunakan sumber daya Anda sendiri, Anda mungkin mendapatkan data dari penyedia data pihak kedua atau pihak ketiga. **Data pihak kedua** dikumpulkan langsung oleh kelompok lain dan kemudian dijual. **Data pihak ketiga** dijual oleh penyedia yang tidak mengumpulkan data itu sendiri. Data pihak ketiga mungkin berasal dari sejumlah sumber berbeda.

**Memecahkan masalah bisnis Anda**

Kumpulan data dapat menampilkan banyak informasi menarik. Namun pastikan untuk memilih data yang benar-benar dapat membantu menyelesaikan pertanyaan masalah Anda. Misalnya, jika Anda menganalisis tren dari waktu ke waktu, pastikan Anda menggunakan data deret waktu — dengan kata lain, data yang menyertakan tanggal.

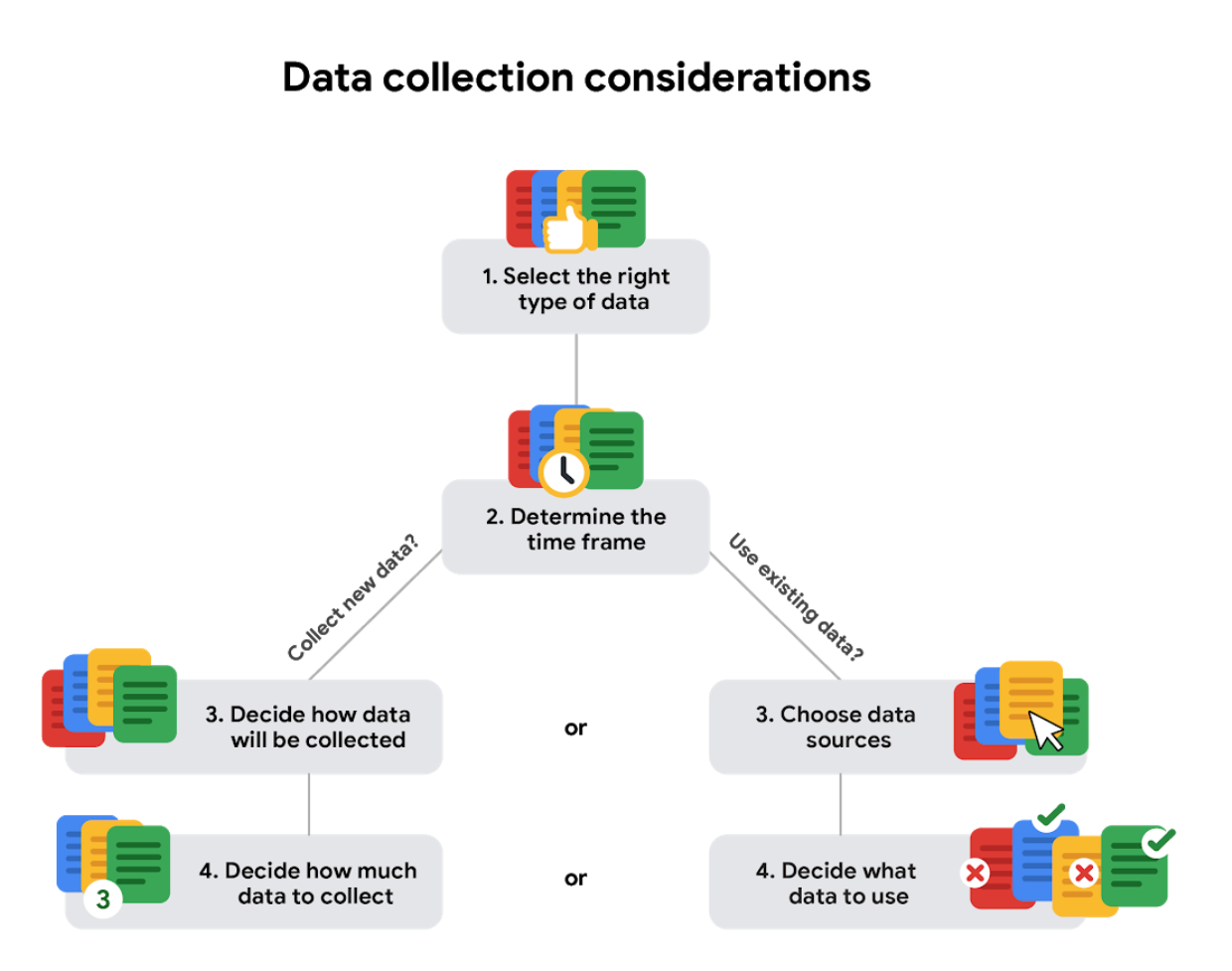
**Berapa banyak data yang harus dikumpulkan**

Jika Anda mengumpulkan data sendiri, buatlah keputusan yang masuk akal mengenai ukuran sampel. Sampel acak dari data yang ada mungkin cocok untuk beberapa proyek. Proyek lain mungkin memerlukan pengumpulan data yang lebih strategis untuk fokus pada kriteria tertentu. Setiap proyek mempunyai kebutuhannya masing-masing.

**Jangka waktu**

Jika Anda mengumpulkan data sendiri, putuskan berapa lama Anda perlu mengumpulkannya, terutama jika Anda melacak tren dalam jangka waktu yang lama. Jika Anda memerlukan jawaban segera, Anda mungkin tidak punya waktu untuk mengumpulkan data baru. Dalam hal ini, Anda perlu menggunakan data historis yang sudah ada.

Gunakan diagram alur di bawah ini jika pengumpulan data sangat bergantung pada berapa banyak waktu yang Anda miliki:



**Format data dalam praktiknya**

Saat Anda memikirkan kata "format", banyak hal yang mungkin terlintas di benak Anda. Bayangkan sebuah iklan untuk toko favorit Anda. Anda mungkin menemukannya dalam bentuk iklan cetak, papan reklame, atau bahkan iklan. Informasi disajikan dalam format yang paling sesuai untuk Anda gunakan. Format kumpulan data sangat mirip dengan itu, dan memilih format yang tepat akan membantu Anda mengelola dan menggunakan data Anda sebaik mungkin.

A computer screen shot of a diagram

Description automatically generated

**Contoh format data**

Seperti kebanyakan hal, definisi akan lebih mudah diklik ketika kita dapat memasangkannya dengan contoh kehidupan nyata. Tinjau setiap definisi terlebih dahulu, lalu gunakan contoh untuk mengunci pemahaman Anda tentang setiap format data.

tabel berikut menyoroti perbedaan antara data primer dan sekunder serta contohnya

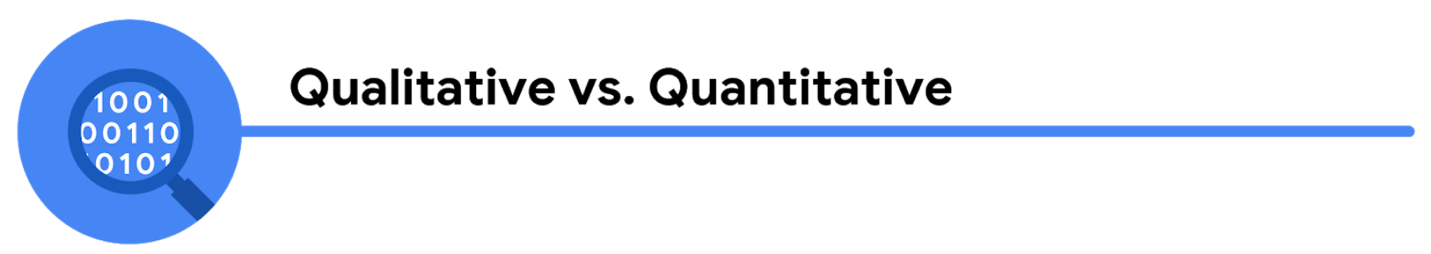
| **Klasifikasi Format Data** | **Definisi** | **Contoh** |
| --- | --- | --- |
| Data utama | Dikumpulkan oleh peneliti dari sumber tangan pertama | - Data dari wawancara yang Anda lakukan - Data dari survei yang diperoleh dari 20 peserta - Data dari kuesioner yang Anda peroleh dari sekelompok pekerja |
| Data sekunder | Dikumpulkan oleh orang lain atau dari penelitian lain | - Data yang Anda beli dari profil pelanggan perusahaan analisis data lokal - Data demografi yang dikumpulkan oleh universitas - Data sensus yang dikumpulkan oleh pemerintah federal |

tabel berikut menyoroti perbedaan antara data internal dan eksternal serta contohnya masing-masing

| **Klasifikasi Format Data** | **Definisi** | **Contoh** |
| --- | --- | --- |
| data dalaman | Data yang ada di dalam sistem perusahaan itu sendiri | - Gaji karyawan di berbagai unit bisnis yang dilacak oleh HR - Data penjualan berdasarkan lokasi toko - Tingkat inventaris produk di seluruh pusat distribusi |
| Data eksternal | Data yang berada di luar perusahaan atau organisasi | - Gaji rata-rata nasional untuk berbagai posisi di seluruh organisasi Anda - Laporan kredit untuk pelanggan dealer mobil |

tabel berikut menyoroti perbedaan antara data kontinu dan data diskrit serta contohnya masing-masing

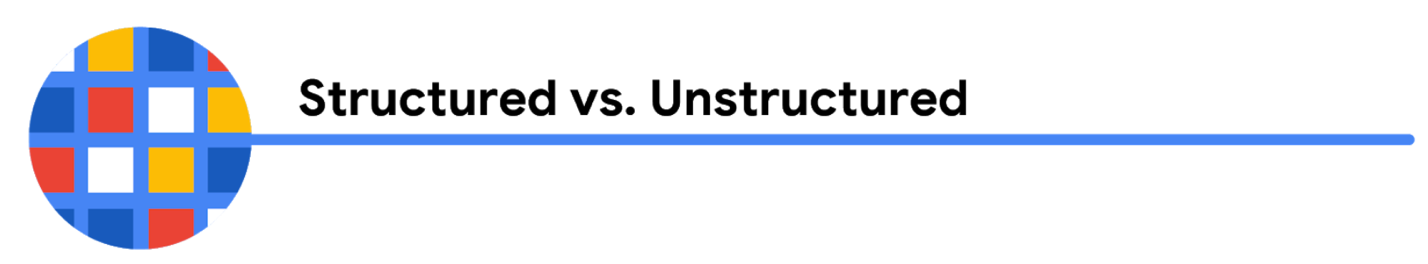
| **Klasifikasi Format Data** | **Definisi** | **Contoh** |
| --- | --- | --- |
| Data berkelanjutan | Data yang diukur dan dapat memiliki hampir semua nilai numerik | - Tinggi badan anak-anak di kelas tiga (52,5 inci, 65,7 inci) - Penanda runtime dalam video - Suhu |
| Data diskrit | Data yang terhitung dan mempunyai jumlah nilai yang terbatas | - Jumlah orang yang mengunjungi rumah sakit setiap hari (10, 20, 200) - Kapasitas maksimum ruangan yang diperbolehkan - Tiket terjual pada bulan berjalan |

tabel berikut menyoroti perbedaan antara data kualitatif dan kuantitatif serta contohnya

| **Klasifikasi Format Data** | **Definisi** | **Contoh** |
| --- | --- | --- |
| Kualitatif | Ukuran kualitas dan karakteristik yang subyektif dan dapat menjelaskan | - Aktivitas olahraga yang paling disukai - Merek favorit dari sebagian besar pelanggan setia - Preferensi fesyen dewasa muda |
| Kuantitatif | Ukuran fakta numerik yang spesifik dan obyektif | - Persentase dokter bersertifikat yang merupakan perempuan - Populasi gajah di Afrika - Jarak dari Bumi ke Mars |

tabel berikut menyoroti perbedaan antara data nominal dan ordinal serta contohnya masing-masing

| **Klasifikasi Format Data** | **Definisi** | **Contoh** |
| --- | --- | --- |
| Nominal | Jenis data kualitatif yang tidak dikategorikan dengan urutan tertentu | - Pelanggan pertama, pelanggan kembali, pelanggan tetap - Pemohon pekerjaan baru, pelamar lama, pelamar internal - Daftar baru, daftar harga dikurangi, penyitaan |
| Urut | Suatu jenis data kualitatif dengan urutan atau skala yang ditetapkan | - Peringkat film (jumlah bintang: 1 bintang, 2 bintang, 3 bintang) - Pilihan pemungutan suara pilihan peringkat (1, 2, 3) - Tingkat pendapatan (penghasilan rendah, pendapatan menengah, pendapatan tinggi) |

tabel berikut menyoroti perbedaan antara data terstruktur dan tidak terstruktur serta contohnya masing-masing

| **Klasifikasi Format Data** | **Definisi** | **Contoh** |
| --- | --- | --- |
| Data terstruktur | Data disusun dalam format tertentu, seperti baris dan kolom | - Laporan pengeluaran - Pengembalian pajak - Menyimpan inventaris |
| Data tidak terstruktur | Data yang tidak disusun dengan cara yang mudah diidentifikasi | - Posting media sosial - Email - Video |

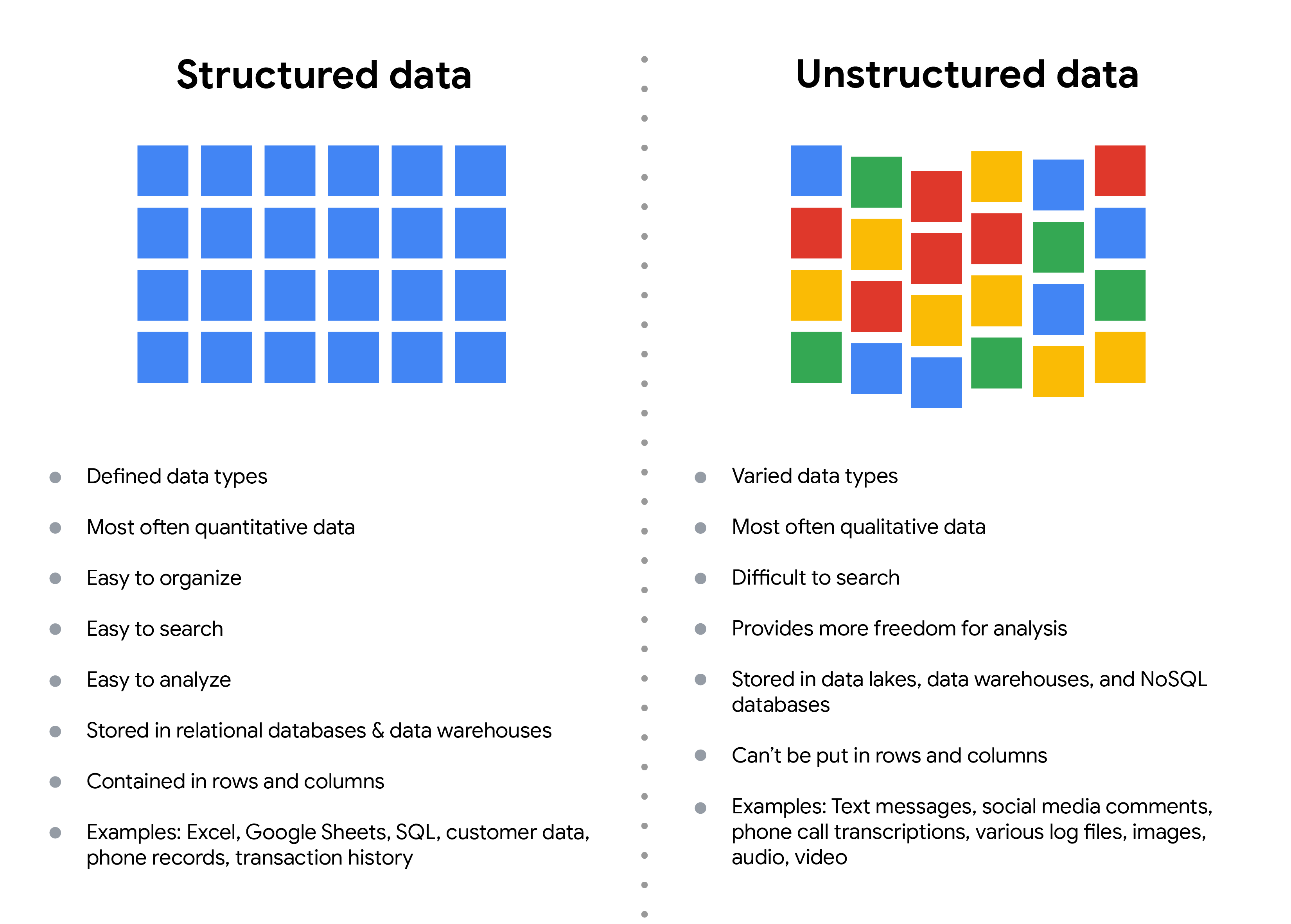
**Struktur data**

Data ada di mana-mana dan dapat disimpan dengan berbagai cara. Dua kategori umum data adalah:

* **Data terstruktur:** Disusun dalam format tertentu, seperti baris dan kolom.
* **Data tidak terstruktur:** Tidak disusun dengan cara yang mudah diidentifikasi.

Misalnya, saat Anda menilai restoran favorit Anda secara online, Anda membuat data terstruktur. Namun saat Anda menggunakan Google Earth untuk memeriksa citra satelit lokasi restoran, Anda menggunakan data tidak terstruktur.

Berikut penyegaran mengenai ciri-ciri data terstruktur dan tidak terstruktur:

Data terstruktur: - Tipe data yang ditentukan - Paling sering data kuantitatif - Mudah diatur - Mudah dicari - Mudah dianalisis - Disimpan dalam database relasional - Termuat dalam baris dan kolom - Contoh: Excel, Google Spreadsheet, SQL, data pelanggan, catatan telepon , riwayat transaksi Data tidak terstruktur: - Tipe data bervariasi - Paling sering data kualitatif - Sulit dicari - Memberikan lebih banyak kebebasan untuk analisis - Disimpan dalam data lake dan database NoSQL - Tidak dapat dimasukkan ke dalam baris dan kolom - Contoh: Pesan teks, sosial komentar media, transkripsi panggilan telepon, berbagai file log, gambar, audio, video

**Data terstruktur**

Seperti yang kami jelaskan sebelumnya, **data terstruktur** disusun dalam format tertentu. Hal ini mempermudah penyimpanan dan kueri kebutuhan bisnis. Jika data diekspor, strukturnya akan sejalan dengan datanya.

**Data tidak terstruktur**

**Data tidak terstruktur** tidak dapat diatur dengan cara yang mudah diidentifikasi. Dan masih banyak lagi data tidak terstruktur dibandingkan data terstruktur di dunia. File video dan audio, file teks, konten media sosial, citra satelit, presentasi, file PDF, tanggapan survei terbuka, dan situs web semuanya memenuhi syarat sebagai jenis data tidak terstruktur.

**Masalah keadilan**

Kurangnya struktur membuat data yang tidak terstruktur sulit dicari, dikelola, dan dianalisis. Namun kemajuan terkini dalam kecerdasan buatan dan algoritma pembelajaran mesin mulai mengubah hal tersebut. Kini, tantangan baru yang dihadapi para ilmuwan data adalah memastikan alat-alat tersebut inklusif dan tidak memihak. Jika tidak, elemen tertentu dari kumpulan data akan lebih berbobot dan/atau terwakili dibandingkan elemen lainnya. Dan saat Anda mempelajarinya, kumpulan data yang tidak adil tidak mewakili populasi secara akurat, sehingga menyebabkan hasil yang tidak merata, tingkat akurasi yang rendah, dan analisis yang tidak dapat diandalkan.

**Tingkat dan teknik pemodelan data**

Bacaan ini memperkenalkan Anda pada pemodelan data dan berbagai jenis model data. Model data membantu menjaga konsistensi data dan memungkinkan orang memetakan cara data diatur. Pemahaman dasar memudahkan analis dan pemangku kepentingan lainnya untuk memahami data mereka dan menggunakannya dengan cara yang benar.

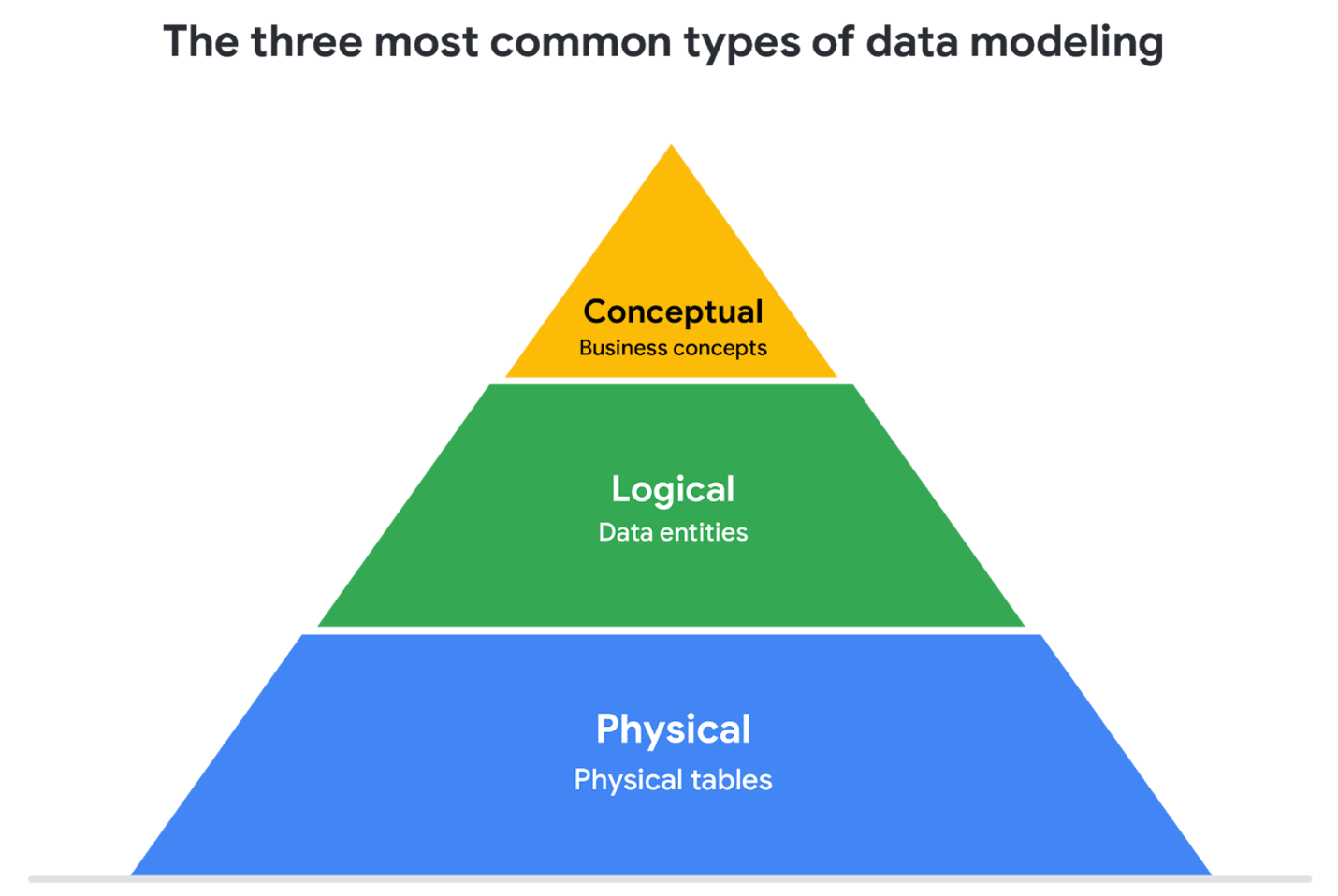
**Catatan penting:** Sebagai analis data junior, Anda tidak akan diminta merancang model data. Namun Anda mungkin menemukan model data yang sudah ada di organisasi Anda.

**Apa itu pemodelan data?**

**Pemodelan data** adalah proses membuat diagram yang secara visual mewakili bagaimana data diatur dan terstruktur. Representasi visual ini disebut **model data** . Anda dapat menganggap pemodelan data sebagai cetak biru sebuah rumah. Kapan saja, mungkin ada tukang listrik, tukang kayu, dan tukang ledeng yang menggunakan cetak biru tersebut. Masing-masing pembangun ini memiliki hubungan yang berbeda dengan cetak birunya, namun mereka semua membutuhkannya untuk memahami struktur rumah secara keseluruhan. Model data serupa; pengguna yang berbeda mungkin memiliki kebutuhan data yang berbeda, namun model data memberi mereka pemahaman tentang struktur secara keseluruhan.

**Tingkat pemodelan data**

Setiap level pemodelan data memiliki tingkat detail yang berbeda-beda.



1. **Pemodelan data konseptual** memberikan gambaran tingkat tinggi tentang struktur data, seperti bagaimana data berinteraksi di seluruh organisasi. Misalnya, model data konseptual dapat digunakan untuk menentukan persyaratan bisnis untuk database baru. Model data konseptual tidak berisi detail teknis.
2. **Pemodelan data logis** berfokus pada rincian teknis database seperti hubungan, atribut, dan entitas. Misalnya, model data logis mendefinisikan bagaimana catatan individual diidentifikasi secara unik dalam database. Tapi itu tidak menyebutkan nama sebenarnya dari tabel database. Itulah tugas model data fisik.
3. **Pemodelan data fisik** menggambarkan bagaimana database beroperasi. Model data fisik mendefinisikan semua entitas dan atribut yang digunakan; misalnya, ini mencakup nama tabel, nama kolom, dan tipe data untuk database.

Informasi lebih lanjut dapat ditemukan di sini[perbandingan model data.](https://www.1keydata.com/datawarehousing/data-modeling-levels.html)

**Teknik pemodelan data**

Ada banyak pendekatan dalam mengembangkan model data, tetapi dua metode yang umum adalah **Entity Relationship Diagram (ERD)** dan diagram **Unified Modeling Language (UML)** . ERD adalah cara visual untuk memahami hubungan antar entitas dalam model data. Diagram UML adalah diagram yang sangat rinci yang menggambarkan struktur suatu sistem dengan menunjukkan entitas sistem, atribut, operasi, dan hubungannya. Sebagai analis data junior, Anda perlu memahami bahwa ada berbagai teknik pemodelan data, namun dalam praktiknya, Anda mungkin akan menggunakan teknik yang sudah ada di organisasi Anda.

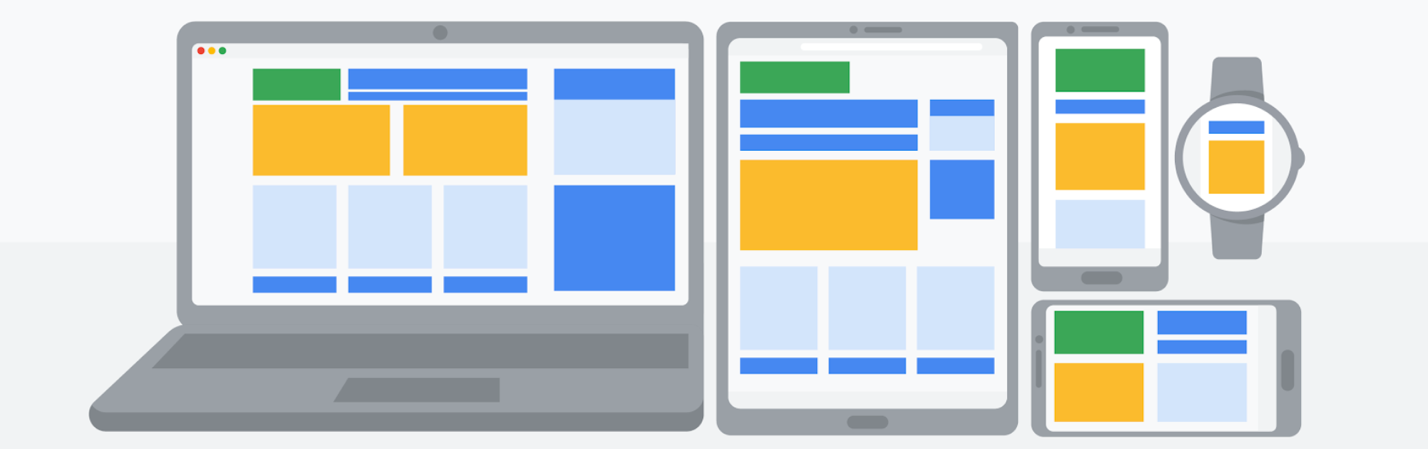
Anda dapat membaca lebih lanjut tentang ERD, UML, dan kamus data di sini[artikel teknik pemodelan data](https://dataedo.com/blog/basic-data-modeling-techniques).

**Analisis data dan pemodelan data**

Pemodelan data dapat membantu Anda menjelajahi detail tingkat tinggi dari data Anda dan bagaimana data tersebut terkait di seluruh sistem informasi organisasi. Pemodelan data terkadang memerlukan analisis data untuk memahami bagaimana data dikumpulkan; dengan begitu, Anda tahu cara memetakan data. Dan yang terakhir, model data memudahkan semua orang di organisasi Anda untuk memahami dan berkolaborasi dengan Anda pada data Anda. Ini penting bagi Anda dan semua orang di tim Anda!

# Memahami logika Boolean

Dalam bacaan ini, Anda akan mempelajari dasar-dasar logika Boolean dan mempelajari cara menggunakan beberapa kondisi dalam pernyataan Boolean. Kondisi ini dibuat dengan operator Boolean, termasuk AND, OR, dan NOT. Operator ini mirip dengan operator matematika dan dapat digunakan untuk membuat pernyataan logis yang memfilter hasil Anda. Analis data menggunakan pernyataan Boolean untuk melakukan berbagai tugas analisis data, seperti membuat kueri untuk penelusuran dan memeriksa kondisi saat menulis kode pemrograman.

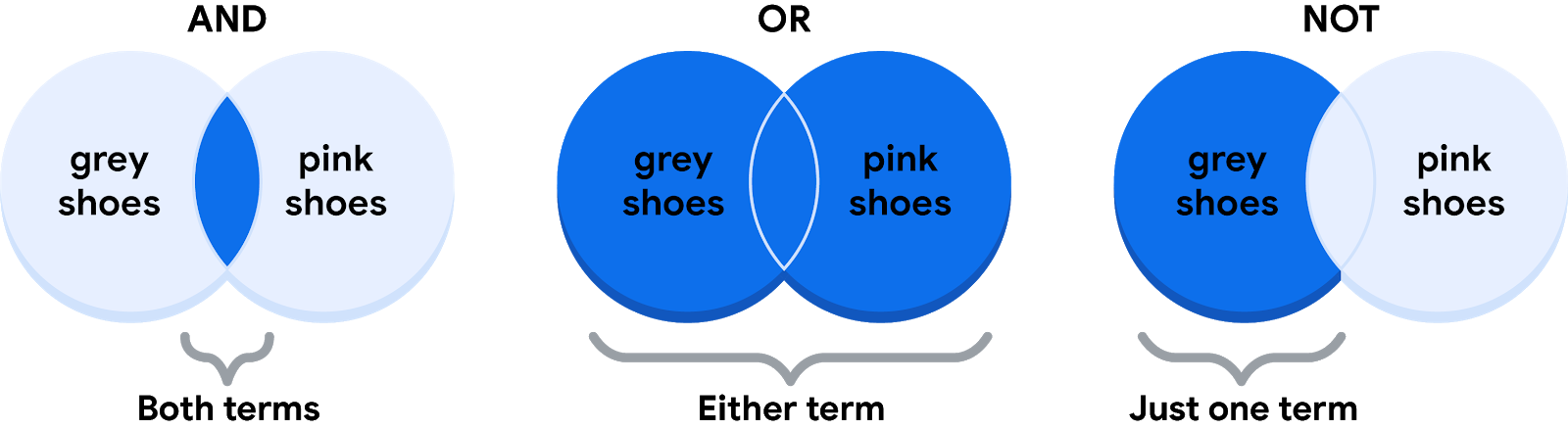


## Contoh logika Boolean

Bayangkan Anda sedang berbelanja sepatu, dan sedang mempertimbangkan preferensi tertentu:

* Anda akan membeli sepatu hanya jika warnanya merah muda dan abu-abu
* Anda akan membeli sepatu jika seluruhnya berwarna merah muda atau abu-abu seluruhnya, atau jika berwarna merah muda dan abu-abu
* Anda akan membeli sepatu jika warnanya abu-abu, tetapi tidak jika warnanya merah jambu

Di bawah ini adalah diagram Venn yang menggambarkan preferensi tersebut. AND adalah pusat diagram Venn, dimana dua kondisi saling tumpang tindih. OR mencakup salah satu kondisi. NOT hanya mencakup bagian diagram Venn yang tidak mengandung pengecualian.



### **Operator DAN**

Syarat Anda adalah “Jika warna sepatu ada kombinasi abu-abu dan merah muda, Anda akan membelinya.” Pernyataan Boolean akan memecah logika pernyataan tersebut untuk memfilter hasil Anda berdasarkan kedua warna. Akan tertulis “JIKA (Warna=”Abu-abu”) DAN (Warna=”Merah Muda”) lalu belilah.” Operator AND memungkinkan Anda menumpuk beberapa kondisi.

Di bawah ini adalah tabel kebenaran sederhana yang menguraikan logika Boolean yang bekerja dalam pernyataan ini. Pada kolom **Color is Grey** terdapat dua pasang sepatu yang memenuhi syarat warna. Dan pada kolom **Color is Pink** terdapat dua pasang yang memenuhi syarat tersebut. Namun pada kolom **If Grey AND Pink** , hanya ada satu pasang sepatu yang memenuhi kedua syarat tersebut. Jadi, menurut logika Boolean dari pernyataan tersebut, hanya ada satu pasangan yang ditandai benar. Dengan kata lain, ada satu pasang sepatu yang bisa Anda beli.

| **Warnanya Abu-abu** | **Warnanya Merah Muda** | **Jika Abu-abu DAN Merah Muda, maka Beli** | **Logika Boolean** |
| --- | --- | --- | --- |
| Abu-abu/Benar | Merah Muda/Benar | Benar/Beli | Benar DAN Benar = Benar |
| Abu-abu/Benar | Hitam/Salah | Salah/Jangan membeli | Benar DAN Salah = Salah |
| Merah/Salah | Merah Muda/Benar | Salah/Jangan membeli | Salah DAN Benar = Salah |
| Merah/Salah | Hijau/Salah | Salah/Jangan membeli | Salah DAN Salah = Salah |

### **Operator ATAU**

Operator OR memungkinkan Anda melanjutkan jika salah satu dari dua kondisi terpenuhi. Syarat Anda adalah “Jika sepatu itu berwarna abu-abu atau merah muda, Anda akan membelinya.” Pernyataan Boolean adalah “JIKA (Warna=”Abu-abu”) ATAU (Warna=”Merah Muda”) maka belilah.” Perhatikan bahwa setiap sepatu yang memenuhi kondisi **Warna Abu-abu** atau **Warna Merah Muda** ditandai sebagai benar oleh logika Boolean. Berdasarkan tabel kebenaran di bawah ini, ada tiga pasang sepatu yang bisa Anda beli.

| **Warnanya Abu-abu** | **Warnanya Merah Muda** | **Jika Abu-abu ATAU Merah Muda, maka Beli** | **Logika Boolean** |
| --- | --- | --- | --- |
| Merah/Salah | Hitam/Salah | Salah/Jangan membeli | Salah ATAU Salah = Salah |
| Hitam/Salah | Merah Muda/Benar | Benar/Beli | Salah ATAU Benar = Benar |
| Abu-abu/Benar | Hijau/Salah | Benar/Beli | Benar ATAU Salah = Benar |
| Abu-abu/Benar | Merah Muda/Benar | Benar/Beli | Benar ATAU Benar = Benar |

### **Operator BUKAN**

Terakhir, operator NOT memungkinkan Anda memfilter dengan mengurangi kondisi tertentu dari hasil. Syarat Anda adalah "Anda akan membeli sepatu abu-abu apa pun kecuali yang ada bekas warna merah jambu di dalamnya." Pernyataan Boolean Anda adalah “JIKA (Warna=”Abu-abu”) DAN (Warna=BUKAN “Merah Muda”) maka belilah.” Sekarang, semua sepatu abu-abu yang bukan merah jambu ditandai benar dengan logika Boolean untuk kondisi **BUKAN Merah Muda** . Sepatu berwarna merah muda ditandai sebagai salah dengan logika Boolean untuk kondisi **BUKAN Merah Muda** . Hanya sepasang sepatu yang dikecualikan dalam tabel kebenaran di bawah ini.

| **Warnanya Abu-abu** | **Warnanya Merah Muda** | **Logika Boolean untuk BUKAN Pink** | **Jika Abu-abu DAN (BUKAN Merah Muda), maka Beli** | **Logika Boolean** |
| --- | --- | --- | --- | --- |
| Abu-abu/Benar | Merah/Salah | Tidak Salah = Benar | Benar/Beli | Benar DAN Benar = Benar |
| Abu-abu/Benar | Hitam/Salah | Tidak Salah = Benar | Benar/Beli | Benar DAN Benar = Benar |
| Abu-abu/Benar | Hijau/Salah | Tidak Salah = Benar | Benar/Beli | Benar DAN Benar = Benar |
| Abu-abu/Benar | Merah Muda/Benar | Tidak Benar = Salah | Salah/Jangan membeli | Benar DAN Salah = Salah |

## Kekuatan berbagai kondisi

Bagi analis data, kekuatan sebenarnya dari logika Boolean berasal dari kemampuan menggabungkan beberapa kondisi dalam satu pernyataan. Misalnya, jika Anda ingin memfilter sepatu yang berwarna abu-abu atau merah muda, dan tahan air, Anda dapat membuat pernyataan Boolean seperti: “IF ((Warna = “Abu-abu”) OR (Warna = “Pink”)) AND (Tahan Air = “Benar”).” Perhatikan bahwa Anda dapat menggunakan tanda kurung untuk mengelompokkan kondisi Anda.

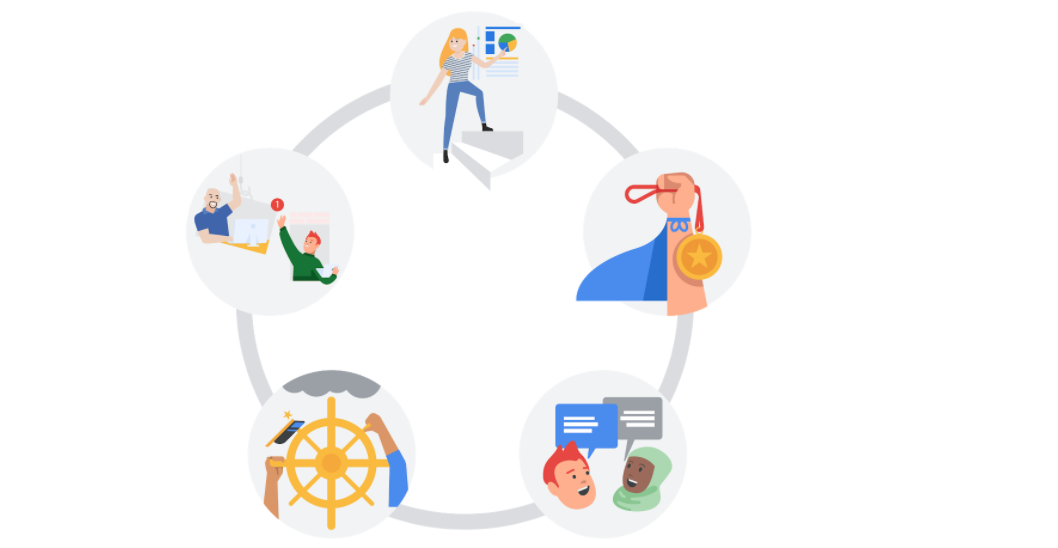
Baik Anda melakukan penelusuran sepatu baru atau menerapkan logika ini ke kueri database Anda, logika Boolean memungkinkan Anda membuat beberapa kondisi untuk memfilter hasil Anda. Dan sekarang setelah Anda mengetahui lebih banyak tentang cara penggunaan logika Boolean, Anda dapat mulai menggunakannya!

## Bacaan/Sumber Tambahan

* Pelajari tentang siapa yang memelopori logika Boolean dalam artikel sejarah ini:[Asal Usul Aljabar Boolean dalam Logika Kelas](https://www.maa.org/press/periodicals/convergence/origins-of-boolean-algebra-in-the-logic-of-classes-george-boole-john-venn-and-c-s-peirce).
* Temukan informasi lebih lanjut tentang penggunaan AND, OR, dan NOT dari ini[tips mencari dengan operator Boolean](https://libguides.mit.edu/c.php?g=175963&p=1158594).

**Mengubah data**

**Apa itu transformasi data?**

Seorang wanita memaparkan data, tangan memegang medali, dua orang mengobrol, kemudi kapal sedang dikemudikan, dua orang saling tos

Dalam bacaan ini, Anda akan mempelajari bagaimana data diubah dan perbedaan antara data lebar dan data panjang. **Transformasi data** adalah proses mengubah format, struktur, atau nilai data. Sebagai seorang analis data, ada kemungkinan besar Anda perlu mengubah data suatu saat nanti untuk memudahkan Anda menganalisisnya.

Transformasi data biasanya melibatkan:

* Menambah, menyalin, atau mereplikasi data
* Menghapus bidang atau catatan
* Standarisasi nama variabel
* Mengganti nama, memindahkan, atau menggabungkan kolom dalam database
* Menggabungkan satu kumpulan data dengan kumpulan data lainnya
* Menyimpan file dalam format berbeda. Misalnya, menyimpan spreadsheet sebagai file nilai yang dipisahkan koma (CSV).

**Mengapa mengubah data?**

Sasaran transformasi data mungkin:

* **Organisasi** data : data yang terorganisir lebih baik lebih mudah digunakan
* **Kompatibilitas** data : aplikasi atau sistem yang berbeda kemudian dapat menggunakan data yang sama
* **Migrasi** data : data dengan format yang cocok dapat dipindahkan dari satu sistem ke sistem lainnya
* **Penggabungan** data : data dengan organisasi yang sama dapat digabungkan menjadi satu
* **Peningkatan** data : data dapat ditampilkan dengan field yang lebih detail
* **Perbandingan** data : perbandingan data yang apel dengan apel kemudian dapat dibuat

**Contoh transformasi data: penggabungan data**

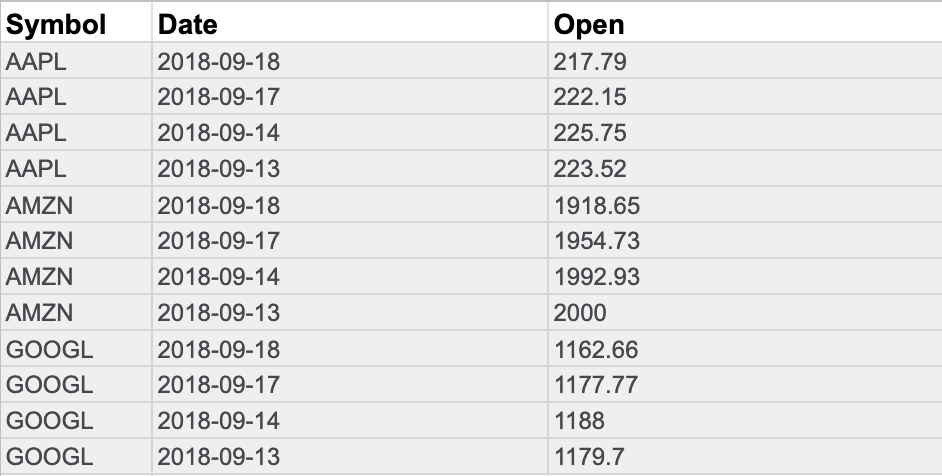
Mario adalah seorang tukang ledeng yang memiliki perusahaan pipa ledeng. Setelah bertahun-tahun berkecimpung dalam bisnis ini, dia membeli perusahaan pipa ledeng lainnya. Mario ingin menggabungkan informasi pelanggan dari perusahaan yang baru diakuisisi dengan miliknya, namun perusahaan lain menggunakan database yang berbeda. Jadi, Mario perlu membuat datanya kompatibel. Untuk melakukan hal ini, ia harus mengubah format data perusahaan yang diakuisisi. Kemudian, dia harus menghapus baris duplikat untuk pelanggan yang sama-sama mereka miliki. Ketika datanya kompatibel dan menyatu, perusahaan pipa Mario akan memiliki database pelanggan yang lengkap dan tergabung.

**Contoh transformasi data: organisasi data (panjang ke lebar)**

Untuk mempermudah pembuatan bagan, Anda mungkin juga perlu mengubah data panjang menjadi data lebar. Perhatikan contoh transformasi harga saham berikut (yang dikumpulkan sebagai data panjang) menjadi data luas.

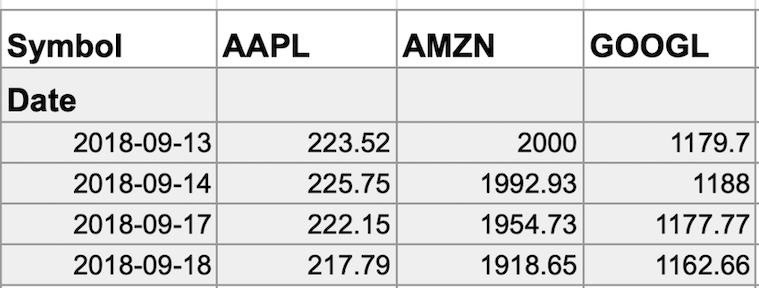
**Data panjang** adalah data yang **setiap barisnya berisi satu titik data** untuk item tertentu. Dalam contoh data panjang di bawah, harga saham individual (titik data) telah dikumpulkan untuk Apple (AAPL), Amazon (AMZN), dan Google (GOOGL) (item tertentu) pada tanggal tertentu.

**Contoh data panjang: Harga saham**



**Data luas** adalah data yang **setiap barisnya berisi beberapa titik data** untuk item tertentu yang diidentifikasi dalam kolom.

**Contoh data luas: Harga saham**



Dengan data yang diubah menjadi data luas, Anda dapat membuat bagan yang membandingkan perubahan saham setiap perusahaan selama periode waktu yang sama.

Anda mungkin memperhatikan bahwa semua data yang disertakan dalam format panjang juga dalam format lebar. Namun data luas lebih mudah dibaca dan dipahami. Itulah sebabnya analis data biasanya lebih sering mengubah data panjang menjadi data lebar daripada mengubah data lebar menjadi data panjang. Tabel berikut merangkum kapan setiap format lebih disukai:

| **Data luas lebih disukai ketika** | **Data yang panjang lebih disukai ketika** |
| --- | --- |
| Membuat tabel dan bagan dengan beberapa variabel tentang setiap mata pelajaran | Menyimpan banyak variabel tentang setiap mata pelajaran. Misalnya, suku bunga senilai 60 tahun untuk setiap bank |
| Membandingkan grafik garis lurus | Melakukan analisis atau grafik statistik tingkat lanjut |