**Basis data dalam analisis data**

Basis data memungkinkan analis untuk memanipulasi, menyimpan, dan memproses data. Ini membantu mereka menelusuri data dengan lebih efisien untuk mendapatkan wawasan terbaik.



**Basis data relasional**

Database **relasional** adalah database yang berisi serangkaian tabel yang dapat dihubungkan untuk menunjukkan hubungan. Pada dasarnya, mereka memungkinkan analis data untuk mengatur dan menghubungkan data berdasarkan kesamaan yang dimiliki data tersebut.

Dalam tabel non-relasional, Anda akan menemukan semua kemungkinan variabel yang mungkin ingin Anda analisis, semuanya dikelompokkan bersama. Hal ini dapat membuat sangat sulit untuk menyelesaikannya. Inilah salah satu alasan mengapa database relasional sangat umum dalam analisis data: database ini menyederhanakan banyak proses analisis dan membuat data lebih mudah ditemukan dan digunakan di seluruh database.

**Normalisasi Basis Data**

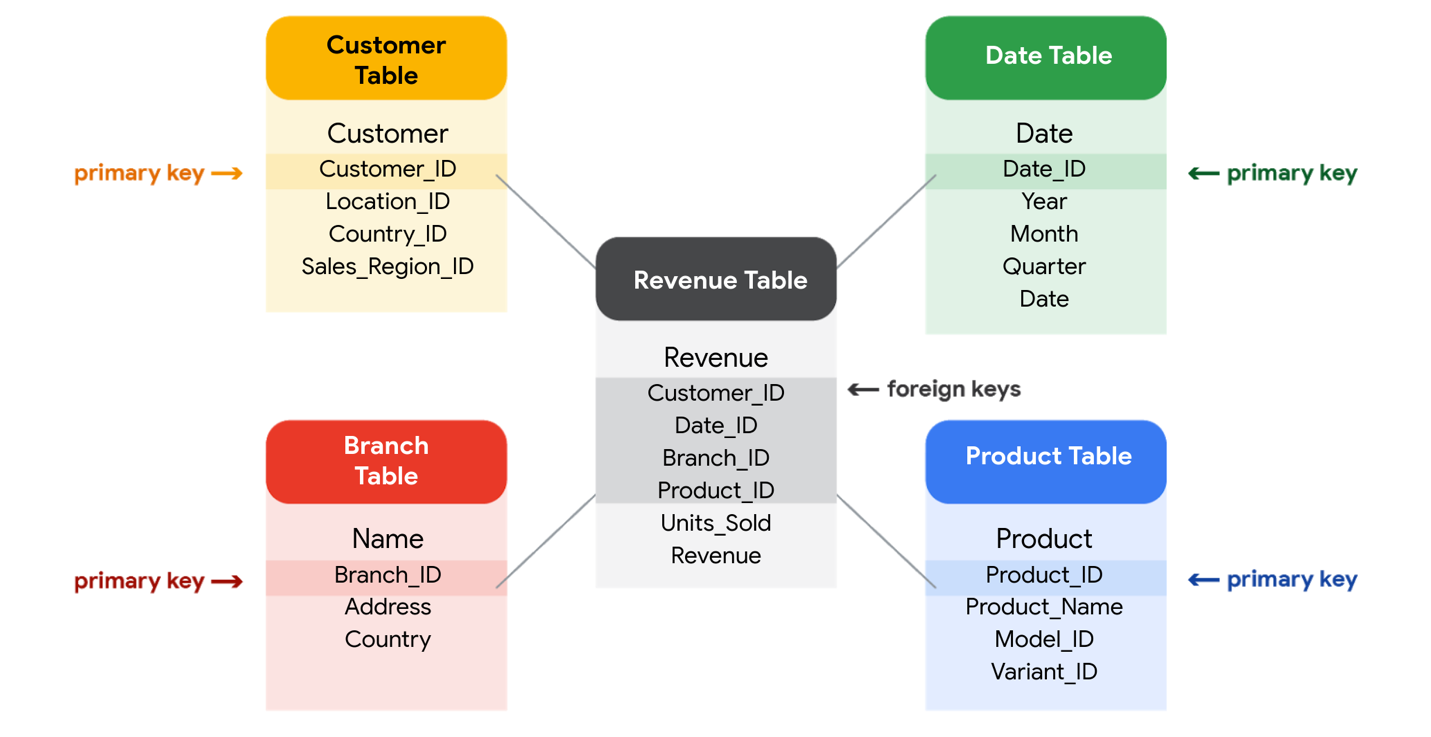
**Normalisasi** adalah proses pengorganisasian data dalam database relasional. Misalnya membuat tabel dan menjalin hubungan antar tabel tersebut. Ini diterapkan untuk menghilangkan redundansi data, meningkatkan integritas data, dan mengurangi kompleksitas dalam database.

**Kunci database relasional**

Tabel-tabel dalam basis data relasional dihubungkan oleh bidang-bidang yang sama yang dimilikinya. Anda mungkin ingat pernah mempelajari tentang kunci primer dan kunci asing sebelumnya. Sebagai pengingat singkat, **kunci utama** adalah pengidentifikasi yang mereferensikan kolom yang setiap nilainya unik. Dengan kata lain, ini adalah kolom tabel yang digunakan untuk mengidentifikasi secara unik setiap rekaman dalam tabel tersebut. Nilai yang ditetapkan ke kunci utama di baris tertentu harus unik di seluruh tabel. Misalnya, jika id\_pelanggan adalah kunci utama untuk tabel pelanggan, tidak akan ada dua pelanggan yang memiliki id\_pelanggan yang sama.

Sebaliknya, **kunci asing** adalah bidang dalam tabel yang merupakan kunci utama dalam tabel lain. Sebuah tabel hanya dapat memiliki satu kunci utama, namun dapat memiliki beberapa kunci asing. Kunci inilah yang membuat hubungan antar tabel dalam database relasional, yang membantu mengatur dan menghubungkan data di beberapa tabel dalam database.

Beberapa tabel tidak memerlukan kunci utama. Misalnya, tabel pendapatan dapat memiliki beberapa kunci asing dan tidak memiliki kunci utama. Kunci utama juga dapat dibuat menggunakan beberapa kolom tabel. Kunci primer jenis ini disebut **kunci komposit** . Misalnya, jika id\_pelanggan dan id\_lokasi adalah dua kolom kunci komposit untuk tabel pelanggan, nilai yang ditetapkan ke bidang tersebut di baris mana pun harus unik dalam keseluruhan tabel.



**SQL? Anda berbicara dalam bahasa saya**

Basis data menggunakan bahasa khusus untuk berkomunikasi yang disebut bahasa query. **Structured Query Language** (SQL) adalah jenis bahasa kueri yang memungkinkan analis data berkomunikasi dengan database. Jadi, seorang analis data akan menggunakan SQL untuk membuat kueri guna melihat data spesifik yang mereka inginkan dari kumpulan yang lebih besar. Dalam database relasional, analis data dapat menulis kueri untuk mendapatkan data dari tabel terkait. SQL adalah alat yang ampuh untuk bekerja dengan database — itulah sebabnya Anda akan mempelajarinya lebih lanjut nanti!

**Memeriksa kumpulan data: Tur langsung dan terpandu**

Sebagai seorang analis data, Anda akan menggunakan data untuk menjawab pertanyaan dan memecahkan masalah. Saat Anda menganalisis data dan menarik kesimpulan, Anda menghasilkan wawasan yang dapat memengaruhi keputusan bisnis, mendorong perubahan positif, dan membantu pemangku kepentingan mencapai tujuan mereka.

Sebelum Anda memulai analisis, penting untuk memeriksa data Anda untuk menentukan apakah data tersebut berisi informasi spesifik yang Anda perlukan untuk menjawab pertanyaan pemangku kepentingan Anda. Dalam kumpulan data mana pun, hal berikut mungkin terjadi:

* Datanya tidak ada (Anda punya data sandwich, tetapi Anda memerlukan data pizza)
* Data tidak mencukupi (Anda memiliki data pizza untuk tanggal 1-7 Juni, namun Anda memerlukan data untuk seluruh bulan Juni)
* Datanya salah (data pizza Anda mencantumkan harga sepotong $250, sehingga Anda mempertanyakan validitas kumpulan data)

Memeriksa kumpulan data akan membantu Anda menentukan pertanyaan mana yang dapat dijawab dan data apa yang masih hilang. Anda mungkin dapat memulihkan data ini dari sumber eksternal atau setidaknya merekomendasikan kepada pemangku kepentingan Anda agar menggunakan sumber data lain.

Dalam bacaan ini, bayangkan Anda seorang analis data yang memeriksa data spreadsheet untuk menentukan apakah mungkin untuk menjawab pertanyaan pemangku kepentingan Anda.

**Skenarionya**

Anda adalah seorang analis data yang bekerja di perusahaan es krim. Manajemen tertarik untuk meningkatkan penjualan es krim perusahaan.

Perusahaan telah mengumpulkan data tentang penjualannya—tetapi tidak banyak. Data yang tersedia berasal dari sumber data internal dan didasarkan pada penjualan tahun 2019. Anda diminta untuk meninjau data dan memberikan beberapa wawasan tentang penjualan es krim perusahaan. Idealnya, manajemen menginginkan jawaban atas pertanyaan-pertanyaan berikut:

1. Apa rasa es krim yang paling populer?
2. Bagaimana suhu mempengaruhi penjualan?
3. Bagaimana akhir pekan dan hari libur mempengaruhi penjualan?
4. Apa perbedaan profitabilitas antara pelanggan baru dan pelanggan lama?

**Unduh datanya**

Anda dapat mengunduh data untuk mengikuti bacaan ini. Untuk menggunakan template data penjualan, klik link di bawah dan pilih “Gunakan Template.”

Tautan ke templat:[Penjualan Es Krim](https://docs.google.com/spreadsheets/d/1NgiKb8wCnJbUTuUkDUiNRpx9NhwncEmoKuPvgfYfOIY/template/preview?resourcekey=0-X3e7NzehG2Y74MIBhOaqeQ#gid=653912415)

ATAU

Jika Anda belum memiliki akun Google, Anda dapat mendownload spreadsheet langsung dari lampiran di bawah:

[PenjualanBerdasarkan Temp](https://d3c33hcgiwev3.cloudfront.net/jmigEulNR7yooBLpTYe8Cw_9ecaf818f1a74b7987fe6a7d9af3c1f1_SalesByTemp.xlsx?Expires=1693785600&Signature=dhLln8w~6OiKLlSqPAKrve2O-js04iRP~dSCl~PeK4Wbj0NJ0TuQ0FrpQKKnYVDpgfeL~5S3UHXCP258-xx6~j0m-64wjxaMd0HJZWZoKoqitnRtoxvEqtKlJyih1~qfQqERpf64QDIw9yADVi--YetXYXIX7GXVKk0hxieEGxg_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A" \t "_blank)

[Berkas XLSX](https://d3c33hcgiwev3.cloudfront.net/jmigEulNR7yooBLpTYe8Cw_9ecaf818f1a74b7987fe6a7d9af3c1f1_SalesByTemp.xlsx?Expires=1693785600&Signature=dhLln8w~6OiKLlSqPAKrve2O-js04iRP~dSCl~PeK4Wbj0NJ0TuQ0FrpQKKnYVDpgfeL~5S3UHXCP258-xx6~j0m-64wjxaMd0HJZWZoKoqitnRtoxvEqtKlJyih1~qfQqERpf64QDIw9yADVi--YetXYXIX7GXVKk0hxieEGxg_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A" \t "_blank)

[Penjualan Berdasarkan Hari](https://d3c33hcgiwev3.cloudfront.net/B3ofmLtERPq6H5i7RFT6Pg_1ca5eec9c08941518e2c16034a2e65f1_SalesByDay.xlsx?Expires=1693785600&Signature=krqueTNFJHVx49DlQaNr71OoVBZUM43kenUVuFAv5t66Z1chIz-i2FhDhmSmTUb5y9UoK57TVDqEUaJ1tbNFdP-sql3W-CDq8VikiUHkCobNWt0i4dauguA2tobcqNltPc9mJ62wpfNmHe5TkcAsirUDKUHpKbKRX7cwwD2Yzaw_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A" \t "_blank)

[Berkas XLSX](https://d3c33hcgiwev3.cloudfront.net/B3ofmLtERPq6H5i7RFT6Pg_1ca5eec9c08941518e2c16034a2e65f1_SalesByDay.xlsx?Expires=1693785600&Signature=krqueTNFJHVx49DlQaNr71OoVBZUM43kenUVuFAv5t66Z1chIz-i2FhDhmSmTUb5y9UoK57TVDqEUaJ1tbNFdP-sql3W-CDq8VikiUHkCobNWt0i4dauguA2tobcqNltPc9mJ62wpfNmHe5TkcAsirUDKUHpKbKRX7cwwD2Yzaw_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A" \t "_blank)

[PenjualanBerdasarkan Rasa](https://d3c33hcgiwev3.cloudfront.net/DHN9hYWCSDCzfYWFgvgwgg_b0e0d35f6a4f4bde9c84ecd0dd69c0f1_SalesByFlavor.xlsx?Expires=1693785600&Signature=XO2vU~oL5Up2sASvhOCJ15SkivFCFH5CaXiCw2IXiQh~wlpcgKsSYfeRKHLMqdngJR8Ogp5okz9D2YPwelgj~Yf29P6q6pjupZBKFBmGOtzrMKfzMUyGXc8XOOrO5AS6JpqcyWvGJtlj3tErfQa8Ln~3DcOB-ngYKJc0GQsfELo_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A" \t "_blank)

[Berkas XLSX](https://d3c33hcgiwev3.cloudfront.net/DHN9hYWCSDCzfYWFgvgwgg_b0e0d35f6a4f4bde9c84ecd0dd69c0f1_SalesByFlavor.xlsx?Expires=1693785600&Signature=XO2vU~oL5Up2sASvhOCJ15SkivFCFH5CaXiCw2IXiQh~wlpcgKsSYfeRKHLMqdngJR8Ogp5okz9D2YPwelgj~Yf29P6q6pjupZBKFBmGOtzrMKfzMUyGXc8XOOrO5AS6JpqcyWvGJtlj3tErfQa8Ln~3DcOB-ngYKJc0GQsfELo_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A" \t "_blank)



**Periksa datanya**

**Pertanyaan 1: Rasa es krim apa yang paling populer?**

Untuk menemukan rasa yang paling populer, pertama-tama Anda perlu mendefinisikan apa yang dimaksud dengan "populer". Apakah rasa yang paling populer merupakan rasa yang menghasilkan pendapatan terbesar di tahun 2019? Ataukah rasa yang paling banyak terjual unitnya di tahun 2019? Terkadang pilihan pengukuran Anda dibatasi oleh data yang Anda miliki—Anda dapat meninjau spreadsheet Anda untuk mengetahui apakah salah satu definisi “populer” ini masuk akal berdasarkan data yang tersedia.

Klik tab **rasa** di spreadsheet Anda untuk melihat data yang relevan. Lembar **rasa** memiliki tiga kolom dan 209 baris data. Header kolomnya adalah **minggu** , **unit terjual** *,* dan **rasa** . Kumpulan data ini tidak dilengkapi dengan deskripsi data, jadi Anda harus mencari tahu sendiri pentingnya kolom tersebut. Berdasarkan data, Anda menyimpulkan bahwa kolom ini memberikan informasi tentang jumlah unit yang terjual untuk setiap rasa es krim, menurut minggu, pada tahun 2019

A screenshot of a spreadsheet

Description automatically generated

Dalam hal ini, Anda dapat mengetahui rasa apa yang paling populer dengan menggunakan unit yang terjual sebagai ukuran Anda. Secara khusus, Anda dapat menggunakan kolom **unit terjual** untuk menghitung jumlah total unit terjual sepanjang tahun untuk setiap rasa *.* Sayangnya, kumpulan data tersebut tidak memberikan jumlah penjualan tahunan berdasarkan rasa. Dalam hal ini, langkah Anda selanjutnya adalah bertanya kepada pemangku kepentingan apakah data penjualan tahunan per rasa tersedia dari sumber lain. Jika tidak, Anda dapat menambahkan pernyataan tentang keterbatasan data saat ini ke dalam analisis Anda.

**Pertanyaan 2: Bagaimana pengaruh suhu terhadap penjualan?**

Untuk menjelajahi pertanyaan kedua Anda, Anda mengklik tab **suhu** dan memeriksa datanya. Lembar **suhu** memiliki dua kolom dan 366 baris data. Tajuk kolomnya adalah **temperatur** dan **penjualan** . Data mungkin menunjukkan total penjualan tahun 2019 per suhu (misalnya, entri pertama mungkin berjumlah $39,69 penjualan selama tiga hari terpisah yang masing-masing memiliki suhu tertinggi 60 derajat). Atau, data mungkin menunjukkan cuplikan penjualan dan suhu setiap hari di tahun 2019 (misalnya, entri pertama mungkin merujuk pada satu hari dengan suhu tertinggi 60 derajat dan penjualan $39,69).

A screenshot of a spreadsheet

Description automatically generated

Jadi yang mana? Ini mungkin cuplikan harian karena ada 365 entri untuk suhu, dan beberapa baris dengan suhu yang sama dan nilai penjualan berbeda. Artinya setiap entri ditujukan untuk satu hari dan bukan ringkasan beberapa hari. Namun, tanpa informasi lebih lanjut, Anda tidak dapat memastikannya. Selain itu, Anda tidak tahu apakah data saat ini dicantumkan dalam urutan berdasarkan tanggal atau dalam urutan berbeda. Langkah Anda selanjutnya adalah menghubungi pemilik kumpulan data untuk klarifikasi.

Jika ternyata suhu memang memengaruhi penjualan, Anda dapat memberikan wawasan seperti berikut kepada pemangku kepentingan: “Ketika suhu tertinggi harian berada di atas X derajat, rata-rata penjualan es krim meningkat sebesar Y. Jadi, bisnis harus merencanakan peningkatan inventaris selama masa-masa ini untuk memaksimalkan penjualan.”

**Pertanyaan 3: Bagaimana pengaruh akhir pekan dan hari libur terhadap penjualan?**

Selanjutnya Anda klik pada tab **penjualan** untuk melihat data tentang tanggal penjualan. Lembar **penjualan** memiliki dua kolom dan 366 baris data. Tajuk kolomnya adalah **tanggal** dan **penjualan** . Data ini kemungkinan besar adalah total penjualan harian di tahun 2019, karena penjualan dicatat setiap tanggal di tahun 2019.

A screenshot of a spreadsheet

Description automatically generated

Anda dapat menggunakan data ini untuk menentukan apakah tanggal tertentu jatuh pada akhir pekan atau hari libur dan menambahkan kolom ke lembar Anda yang mencerminkan informasi ini. Kemudian, Anda bisa mengetahui apakah penjualan di akhir pekan dan hari libur lebih besar dibandingkan penjualan di hari lainnya. Hal ini berguna untuk diketahui untuk perencanaan inventaris dan tujuan pemasaran.

**Pertanyaan 4: Apa perbedaan profitabilitas antara pelanggan baru dan pelanggan lama?**

Dataset Anda tidak berisi data penjualan terkait pelanggan baru. Tanpa data ini, Anda tidak akan bisa menjawab pertanyaan terakhir Anda. Namun, mungkin saja perusahaan mengumpulkan data pelanggan dan menyimpannya di tabel data yang berbeda.

Jika ya, langkah Anda selanjutnya adalah mencari tahu cara mengakses data pelanggan perusahaan. Anda kemudian dapat menggabungkan data pendapatan penjualan ke tabel data pelanggan untuk mengkategorikan setiap penjualan berdasarkan pelanggan baru atau pelanggan kembali dan menganalisis perbedaan profitabilitas antara dua kelompok pelanggan. Informasi ini akan membantu pemangku kepentingan Anda mengembangkan kampanye pemasaran untuk jenis pelanggan tertentu guna meningkatkan loyalitas merek dan profitabilitas secara keseluruhan.

**Kesimpulan**

Saat mengerjakan proyek analitik, Anda tidak selalu memiliki semua data yang diperlukan atau relevan. Dalam banyak kasus, Anda dapat beralih ke sumber data lain untuk mengisi kekosongan tersebut.

 Meskipun kumpulan data Anda terbatas, Anda masih bisa memberikan wawasan berharga kepada pemangku kepentingan. Untuk langkah selanjutnya, rencana tindakan terbaik Anda adalah mengambil inisiatif untuk mengajukan pertanyaan, mengidentifikasi kumpulan data lain yang relevan, atau melakukan riset sendiri. Apa pun data yang Anda gunakan, memeriksa data dengan cermat akan memberikan dampak besar pada kualitas analisis Anda secara keseluruhan.

**Metadata sama pentingnya dengan data itu sendiri**

Analisis data, menurut desainnya, adalah bidang yang berkembang pesat dalam pengumpulan dan pengorganisasian data. Dalam bacaan ini, Anda akan belajar tentang cara menganalisis dan memahami secara menyeluruh setiap aspek data Anda.

A group of people standing next to a website

Description automatically generated

Lihatlah data apa pun yang Anda temukan. Apa itu? Dari mana asalnya? Apakah itu berguna? Bagaimana Anda tahu? Di sinilah metadata berperan untuk memberikan pemahaman yang lebih mendalam tentang data. Sederhananya, **metadata** adalah data tentang data. Dalam manajemen basis data, ini memberikan informasi tentang data lain dan membantu analis data menafsirkan konten data dalam basis data.

Terlepas dari apakah Anda bekerja dengan data dalam jumlah besar atau kecil, metadata adalah ciri tim analitik yang berpengetahuan luas, membantu mengomunikasikan data di seluruh bisnis dan mempermudah penggunaan kembali data. Intinya, metadata memberitahukan siapa, apa, kapan, di mana, yang mana, bagaimana, dan mengapa suatu data.

**Elemen metadata**

Sebelum melihat contoh metadata, penting untuk memahami jenis informasi apa yang biasanya disediakan metadata.

**Judul dan deskripsi**

Apa nama file atau situs web yang Anda periksa? Jenis konten apa yang dikandungnya?

**Tag dan kategori**

Apa gambaran umum data yang Anda miliki? Apakah data diindeks atau dijelaskan dengan cara tertentu?

**Siapa yang menciptakannya dan kapan**

Dari mana data tersebut berasal dan kapan data tersebut dibuat? Apakah baru atau sudah ada sejak lama?

**Siapa terakhir kali memodifikasinya dan kapan**

Apakah ada perubahan yang dilakukan pada data? Jika ya, apakah modifikasinya terkini?

**Siapa yang dapat mengakses atau memperbaruinya**

Apakah kumpulan data ini bersifat publik? Apakah izin khusus diperlukan untuk menyesuaikan atau mengubah kumpulan data?

**Contoh metadata**

Di dunia digital saat ini, metadata ada di mana-mana, dan menyediakan metadata di banyak media dan informasi yang berinteraksi dengan Anda sudah menjadi praktik yang lebih umum. Berikut beberapa contoh nyata tempat menemukan metadata:

**Foto**

Setiap kali foto diambil dengan kamera, metadata seperti nama file kamera, tanggal, waktu, dan geolokasi dikumpulkan dan disimpan bersamanya.

**Email**

Saat email dikirim atau diterima, ada banyak metadata yang terlihat seperti baris subjek, pengirim, penerima, serta tanggal dan waktu pengiriman. Ada juga metadata tersembunyi yang mencakup nama server, alamat IP, format HTML, dan detail perangkat lunak.

**Spreadsheet dan dokumen**

Spreadsheet dan dokumen sudah diisi dengan sejumlah besar data sehingga tidak mengherankan jika metadata juga menyertainya. Judul, penulis, tanggal pembuatan, jumlah halaman, komentar pengguna serta nama tab, tabel, dan kolom adalah metadata yang dapat ditemukan di spreadsheet dan dokumen.

**Situs web**

Setiap halaman web memiliki sejumlah bidang metadata standar, seperti tag dan kategori, nama pembuat situs, judul dan deskripsi halaman web, waktu pembuatan, dan ikonografi apa pun.

**File digital**

Biasanya, jika Anda mengklik kanan pada file komputer mana pun, Anda akan melihat metadatanya. Ini dapat terdiri dari nama file, ukuran file, tanggal pembuatan dan modifikasi, dan jenis file.

**Buku**

Metadata tidak hanya digital. Setiap buku memiliki sejumlah metadata standar di sampul dan di dalamnya yang akan memberi tahu Anda tentang judulnya, nama penulis, daftar isi, informasi penerbit, uraian hak cipta, indeks, dan uraian singkat isi buku.

**Data seperti yang Anda tahu**

Mengetahui konten dan konteks data Anda, serta strukturnya, sangat berharga dalam karier Anda sebagai analis data. Saat menganalisis data, penting untuk selalu memahami gambaran lengkapnya. Ini bukan hanya tentang data yang Anda lihat, namun bagaimana data tersebut digabungkan. Metadata memastikan bahwa Anda dapat menemukan, menggunakan, menyimpan, dan menggunakan kembali data di masa mendatang. Ingat, Anda bertanggung jawab untuk mengelola dan menggunakan data secara keseluruhan; metadata sama pentingnya dengan data itu sendiri.

**From external source to a spreadsheet**

When you work with spreadsheets, there are a few different ways to import data. This reading covers how you can import data from external sources, specifically:

* Other spreadsheets
* CSV files
* HTML tables (in web pages)

**Importing data from other spreadsheets**

In a lot of cases, you might have an existing spreadsheet open and need to add additional data from another spreadsheet.

**Google Sheets**

In Google Sheets, you can use the IMPORTRANGE function. It enables you to specify a range of cells in the other spreadsheet to duplicate in the spreadsheet you are working in. You must allow access to the spreadsheet containing the data the first time you import the data. **The URL shown below is for syntax purposes only. Don't enter it in your own spreadsheet. Replace it with a URL to a spreadsheet you have created so you can control access to it by clicking the Allow access button.**

A screenshot of a spreadsheet

Description automatically generated

Refer to the Google Help Center's [IMPORTRANGE](https://support.google.com/docs/answer/3093340?hl=en&ref_topic=9199554) page for more information about the syntax. There is also an example of its use later in the program in [Advanced functions for speedy data cleaning](https://www.coursera.org/learn/process-data/supplement/PLnRS/advanced-functions-for-speedy-data-cleaning).

**Microsoft Excel**

To import data from another spreadsheet, do the following:

**Step 1:** Select **Data** from the main menu.

**Step 2:** Click **Get Data**, and then select **From File** within the toolbar. In the drop down, choose **From Excel Workbook**

**Step 3:** Browse for and select the spreadsheet file and then click **Import**.

**Step 4:** In the Navigator, select which worksheet to import.

**Step 5:** Click **Load** to import all the data in the worksheet; or click **Transform Data** to open the Power Query Editor to adjust the columns and rows of data you want to import.

**Step 6:** If you clicked Transform Data, click **Close & Load** and then select one of the two options:

* **Close & Load** - import the data to a new worksheet
* **Close & Load to...** - import the data to an existing worksheet

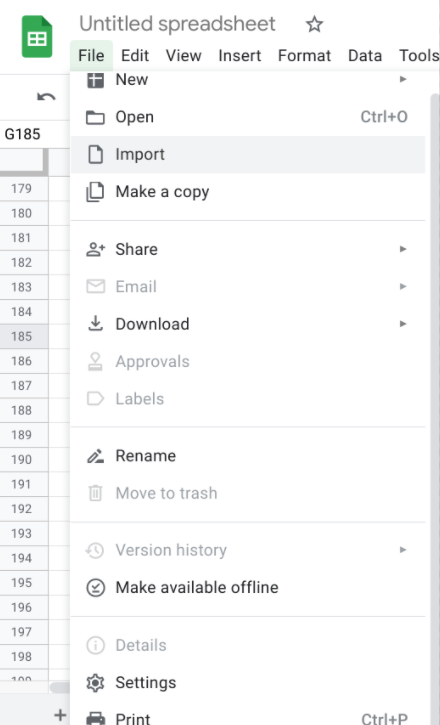
If these directions do not work for the version of Excel that you have. Visit this free online training center, [Microsoft Excel for Windows Training](https://support.microsoft.com/en-us/office/excel-video-training-9bc05390-e94c-46af-a5b3-d7c22f6990bb), you will find everything you need to know, all in one place.

If you are using Numbers, search the [Numbers User Guide](https://support.apple.com/guide/numbers/welcome/mac) for directions.

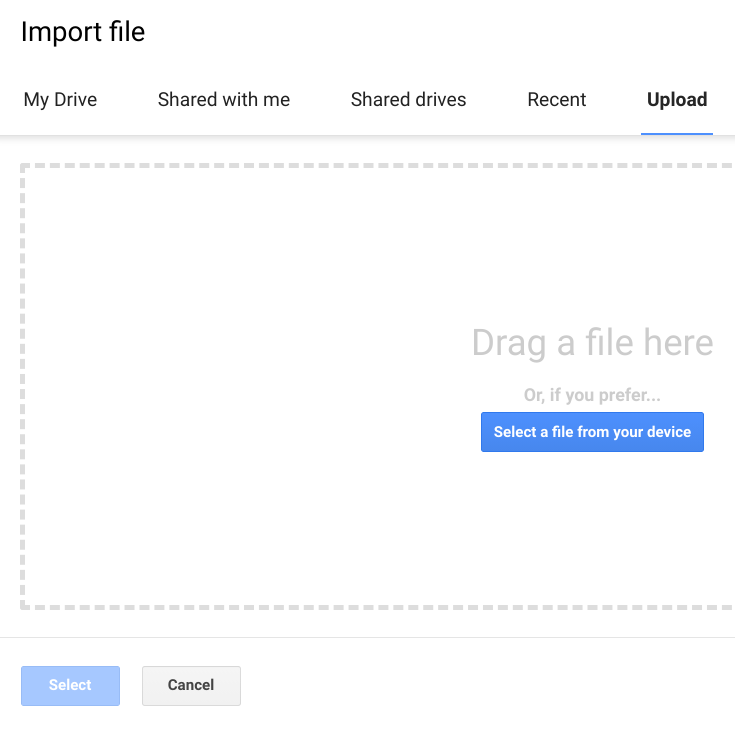
**Importing data from CSV files**

**Google Sheets**

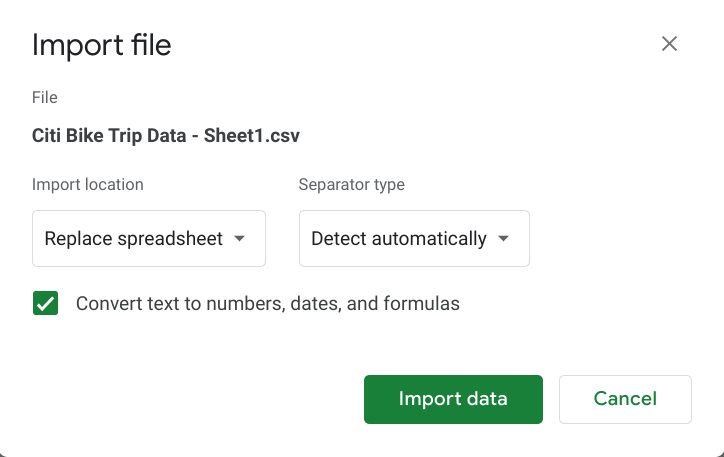
**Step 1:** Open the **File** menu in your spreadsheet and select **Import** to open the Import file window.



**Step 2:** Select **Upload** andthenselect the CSV file you want to import.



**Step 3:** From here, you will have a few options. For **Import location**, you can choose to replace the current spreadsheet, create a new spreadsheet, insert the CSV data as a new sheet, add the data to the current spreadsheet, or replace the data in a specific cell. The data will be inserted as plain text only if you *uncheck* the Convert text to numbers, dates, and formulas checkbox, which is the default setting. Sometimes a CSV file uses a separator like a semi-colon or even a blank space instead of a comma. For **Separator type**, you can select Tab or Comma, or select Custom to enter another character that is being used as the separator.



**Step 4:** Select **Import data**. The data in the CSV file will be loaded into your sheet, and you can begin using it!

**Note:** You can also use the **IMPORTDATA** function in a spreadsheet cell to import data using the URL to a CSV file. Refer to Google Help Center's [IMPORTDATA](https://support.google.com/docs/answer/3093335?hl=en) page for more information and the syntax.

**Microsoft Excel**

**Step 1:** Open a new or existing spreadsheet

**Step 2:** Click **Data** in the main menu and select the **From Text/CSV** option.

**Step 3:** Browse for and select the CSV file and then click **Import**.

**Step 4:** From here, you will have a few options. You can change the delimiter from a comma to another character such as a semicolon. You can also turn automatic data type detection on or off. And, finally, you can transform your data by clicking **Transform Data** to open the Power Query Editor.

**Step 5:** In most cases, accept the default settings in the previous step and click **Load** to load the data in the CSV file to the spreadsheet. The data in the CSV file will be loaded into the spreadsheet, and you can begin working with the data.

If these directions do not work for the version of Excel that you have. Visit this free online training center, [Microsoft Excel for Windows Training](https://support.microsoft.com/en-us/office/excel-video-training-9bc05390-e94c-46af-a5b3-d7c22f6990bb), you will find everything you need to know, all in one place.

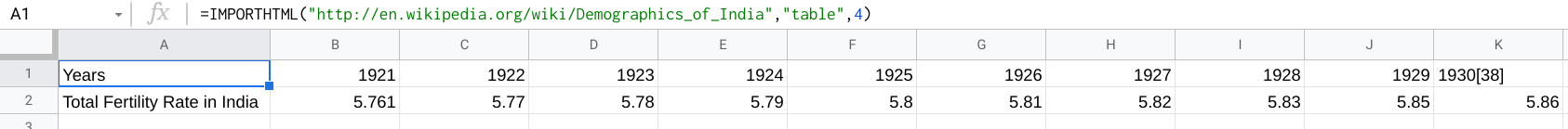
If you are using Numbers, search the [Numbers User Guide](https://support.apple.com/guide/numbers/welcome/mac) for directions.

**Importing HTML tables from web pages**

Importing HTML tables is a very basic method to extract or "scrape" data from public web pages. [Web scraping made easy](https://www.thedataschool.co.uk/anna-prosvetova/web-scraping-made-easy-import-html-tables-or-lists-using-google-sheets-and-excel) introduces how to do this with Google Sheets or Microsoft Excel.

**Google Sheets**

In Google Sheets, you can use the **IMPORTHTML** function. It enables you to import the data from an HTML table (or list) on a web page.



Refer to the Google Help Center's [IMPORTHTML](https://support.google.com/docs/answer/3093339?hl=en) page for more information about the syntax. If you are importing a list, replace "table" with "list" in the above example. The number 4 is the index that refers to the order of the tables on a web page. It is like a pointer indicating which table on the page you want to import the data from.

***You can try this yourself!*** In blank worksheets, copy and paste each of the following IMPORTHTML functions into cell A1 and watch what happens. You will actually be importing the data from four different HTML tables in a Wikipedia article: [Demographics of India](https://en.wikipedia.org/wiki/Demographics_of_India). You can compare your imported data with the tables in the article.

* =IMPORTHTML("http://en.wikipedia.org/wiki/Demographics\_of\_India","table",1)
* =IMPORTHTML("http://en.wikipedia.org/wiki/Demographics\_of\_India","table",2)
* =IMPORTHTML("http://en.wikipedia.org/wiki/Demographics\_of\_India","table",3)
* =IMPORTHTML("http://en.wikipedia.org/wiki/Demographics\_of\_India","table",4)

**Microsoft Excel**

You can import data from web pages using the **From Web** option:

**Step 1:** Open a new or existing spreadsheet.

**Step 2:** Click Data in the main menu and select the **From Web** option.

**Step 3:** Enter the URL and click OK.

**Step 4:** In the Navigator, select which table to import.

**Step 5:** Click **Load** to load the data from the table into your spreadsheet.

If these directions do not work for the version of Excel that you have. Visit this free online training center, [Microsoft Excel for Windows Training](https://support.microsoft.com/en-us/office/excel-video-training-9bc05390-e94c-46af-a5b3-d7c22f6990bb), you will find everything you need to know, all in one place.

If you are using Numbers, search the [Numbers User Guide](https://support.apple.com/guide/numbers/welcome/mac) for directions

**Menjelajahi kumpulan data publik**

**Data terbuka** membantu membuat banyak **kumpulan data publik** yang dapat Anda akses untuk membuat keputusan berdasarkan data. Berikut beberapa sumber daya yang dapat Anda gunakan untuk mulai mencari sendiri kumpulan data publik:

* Itu[Kumpulan Data Publik Google Cloud](https://cloud.google.com/public-datasets)memungkinkan analis data mengakses kumpulan data publik dengan permintaan tinggi, dan mempermudah mengungkap wawasan di cloud.
* Itu[Pencarian Kumpulan Data](https://datasetsearch.research.google.com/)dapat membantu Anda menemukan kumpulan data yang tersedia secara online dengan pencarian kata kunci.
* [Kaggle](https://www.kaggle.com/datasets?utm_medium=paid&utm_source=google.com+search&utm_campaign=datasets&gclid=CjwKCAiAt9z-BRBCEiwA_bWv-L6PpACh6RzmrJjQjmNGCCE7kky1FCtc6Jf1qld-4NwDMYL0WsUyxBoCdwAQAvD_BwE)memiliki fungsi pencarian Data Terbuka yang dapat membantu Anda menemukan kumpulan data untuk berlatih.
* Akhirnya,[Kueri Besar](https://cloud.google.com/bigquery/public-data)menampung 150+ kumpulan data publik yang dapat Anda akses dan gunakan.

**Kumpulan data kesehatan masyarakat**

1. [Data Observatorium Kesehatan Global](https://www.who.int/data/collections): Anda dapat mencari kumpulan data dari halaman ini atau menjelajahi koleksi data unggulan dari Organisasi Kesehatan Dunia.
2. [Kumpulan data Arsip Pencitraan Kanker (TCIA).](https://cloud.google.com/healthcare/docs/resources/public-datasets/tcia): Sama seperti kumpulan data sebelumnya, data ini dihosting oleh Kumpulan Data Publik Google Cloud dan dapat diunggah ke BigQuery.
3. [1000 Genom](https://cloud.google.com/life-sciences/docs/resources/public-datasets/1000-genomes): Ini adalah kumpulan data lain dari sumber daya Publik Google Cloud yang dapat diunggah ke BigQuery.

**Kumpulan data iklim publik**

1. [Pusat Data Iklim Nasional](https://www.ncei.noaa.gov/products): Halaman Tautan Cepat NCDC memiliki pilihan kumpulan data yang dapat Anda jelajahi.
2. [Galeri Kumpulan Data Publik NOAA](https://www.climate.gov/maps-data/datasets): Galeri Kumpulan Data Publik NOAA berisi kumpulan kumpulan data publik yang dapat dicari.

**Kumpulan data sosial-politik publik**

1. [UNICEF Negara Anak-anak Sedunia](https://data.unicef.org/resources/dataset/sowc-2019-statistical-tables/): Dataset dari UNICEF ini mencakup kumpulan tabel yang dapat diunduh.
2. [Statistik Angkatan Kerja CPS](https://www.bls.gov/cps/tables.htm): Halaman ini berisi link ke beberapa dataset yang tersedia yang dapat Anda jelajahi.
3. [Proyek Perpolisian Terbuka Stanford](https://openpolicing.stanford.edu/): Kumpulan data ini dapat diunduh sebagai file .CSV untuk Anda gunakan sendiri.

**Menggunakan BigQuery**

[Kueri Besar](https://cloud.google.com/bigquery/docs)adalah gudang data di Google Cloud yang dapat digunakan analis data untuk membuat kueri, memfilter kumpulan data besar, mengagregasi hasil, dan melakukan operasi kompleks.

Aktivitas mendatang dilakukan di BigQuery. Bacaan ini memberikan petunjuk untuk membuat akun BigQuery Anda sendiri, memilih set data publik, dan mengupload file CSV. Di akhir pembacaan ini, Anda dapat mengonfirmasi akses ke konsol BigQuery sebelum melanjutkan ke aktivitas,

**Catatan:** Sumber daya tambahan untuk memulai untuk beberapa platform database SQL lainnya juga disediakan di akhir bacaan ini jika Anda memilih untuk menggunakannya daripada BigQuery.

**Jenis akun BigQuery**

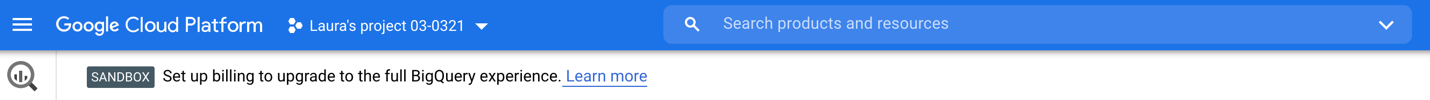
Ada dua jenis akun yang berbeda: sandbox dan uji coba gratis. Akun sandbox memungkinkan Anda mempraktikkan kueri dan menjelajahi kumpulan data publik secara gratis, tetapi memiliki tambahan[pembatasan](https://cloud.google.com/bigquery/docs/sandbox#limits)di atas[kuota dan batasan standar](https://cloud.google.com/bigquery/quotas). Jika Anda lebih suka menggunakan BigQuery dengan batasan standar, Anda dapat menyiapkan akun uji coba gratis. Keterangan lebih lanjut:

* Akun **sandbox gratis** tidak menanyakan metode pembayaran. Namun, ini membatasi Anda hingga 12 proyek. Itu juga tidak memungkinkan Anda memasukkan catatan baru ke database atau memperbarui nilai bidang catatan yang sudah ada. Operasi bahasa manipulasi data (DML) ini tidak didukung di sandbox.
* Akun **uji coba gratis** memerlukan metode pembayaran untuk membuat akun yang dapat ditagih, namun menawarkan fungsionalitas penuh selama masa uji coba.

Dengan jenis akun mana pun, Anda dapat meningkatkan ke akun berbayar kapan saja dan mempertahankan semua proyek yang ada. Jika Anda menyiapkan akun uji coba gratis tetapi memilih untuk tidak meningkatkan ke akun berbayar saat masa uji coba berakhir, Anda masih dapat menyiapkan akun sandbox gratis pada saat itu. Namun, proyek dari akun uji coba Anda tidak akan ditransfer ke akun sandbox Anda. Ini seperti memulai dari awal lagi.

**Siapkan akun sandbox gratis untuk digunakan dalam program ini**

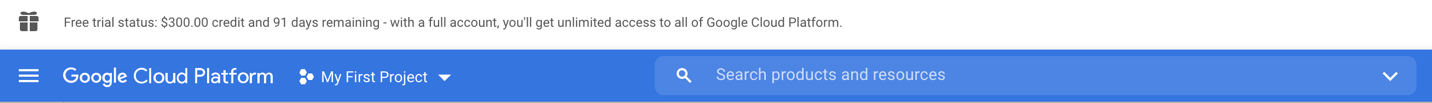
* Ikuti ini[petunjuk langkah demi langkah](https://cursive.io/shared/2da0e63f3-9de7-476f-997b-93fff70d7cb6)atau tonton videonya,[Menyiapkan BigQuery, termasuk opsi sandbox dan penagihan](https://www.coursera.org/learn/data-preparation/lecture/YCkys/setting-up-bigquery-including-sandbox-and-billing-options).
* Untuk informasi lebih detail tentang penggunaan sandbox, mulailah dengan dokumentasi,[Menggunakan kotak pasir BigQuery](https://cloud.google.com/bigquery/docs/sandbox?hl=en_US).
* Setelah menyiapkan akun, Anda akan melihat nama proyek yang Anda buat untuk akun tersebut di banner dan **SANDBOX** di bagian atas konsol BigQuery.



**Siapkan akun uji coba gratis (jika Anda mau)**

Jika Anda memilih untuk tidak memiliki batasan sandbox di BigQuery, Anda dapat menyiapkan akun uji coba gratis untuk digunakan dalam program ini.

* Ikuti ini[petunjuk langkah demi langkah](https://cursive.io/shared/2e98bf922-42d6-48c2-998f-6057389d0ccb)atau tonton videonya,[Menyiapkan BigQuery, termasuk opsi sandbox dan penagihan](https://www.coursera.org/learn/data-preparation/lecture/YCkys/setting-up-bigquery-including-sandbox-and-billing-options). Uji coba gratis menawarkan kredit $300 selama 90 hari ke depan. Anda tidak akan mendekati batas pembelanjaan tersebut jika Anda hanya menggunakan konsol BigQuery untuk mempraktikkan kueri SQL. Setelah Anda membelanjakan kredit $300 (atau setelah 90 hari), uji coba gratis Anda akan berakhir dan Anda harus memilih secara pribadi untuk meningkatkan ke akun berbayar agar tetap menggunakan layanan Google Cloud Platform, termasuk BigQuery. **Metode pembayaran Anda tidak akan ditagih secara otomatis setelah uji coba gratis Anda berakhir.** Jika Anda memilih untuk meningkatkan akun Anda, Anda akan mulai ditagih.
* Setelah Anda menyiapkan akun, Anda akan melihat **Proyek Pertama Saya** di spanduk dan status akun Anda di atas spanduk **–** saldo kredit Anda dan jumlah hari yang tersisa dalam masa percobaan Anda.



**Cara membuka konsol BigQuery**

Di browser Anda, buka[console.cloud.google.com/bigquery](https://console.cloud.google.com/bigquery).

**Catatan:** Pergi ke[konsol.cloud.google.com](https://console.cloud.google.com/)di browser Anda akan membawa Anda ke dasbor utama Google Cloud Platform. Untuk membuka BigQuery dari dasbor, lakukan hal berikut:

* Klik ikon menu Navigasi (ikon Hamburger) di banner.
* Gulir ke bawah ke bagian **DATA BESAR** .
* Klik **BigQuery** dan pilih **ruang kerja SQL** .

Perhatikan[Cara menggunakan BigQuery](https://www.coursera.org/learn/data-preparation/lecture/YWn81/how-to-use-bigquery)video untuk pengenalan setiap bagian ruang kerja BigQuery SQL.

**(Opsional) Jelajahi kumpulan data publik BigQuery**

Anda akan menjelajahi kumpulan data publik dalam aktivitas mendatang, sehingga Anda dapat melakukan langkah-langkah ini nanti jika Anda mau.

* Lihat ini[petunjuk langkah demi langkah](https://cursive.io/shared/242bde9a6-5415-4ce0-bbae-7e875d14d927).

**(Opsional) Unggah file CSV ke BigQuery**

Langkah-langkah ini disediakan agar Anda dapat mengerjakan sendiri kumpulan data saat ini. Anda akan mengupload file CSV ke BigQuery nanti di program.

* Lihat ini [petunjuk langkah demi langkah](https://cursive.io/shared/2dea0d610-ef6b-4ba8-8e44-d40dfeb0454b).

**Memulai database lain (jika tidak menggunakan BigQuery)**

Aktivitas kursus akan lebih mudah diikuti jika Anda menggunakan BigQuery, namun jika Anda terhubung ke dan mempraktikkan kueri SQL di platform database lain selain BigQuery, berikut sumber daya serupa untuk memulai:

* [Memulai MySQL](https://dev.mysql.com/doc/mysql-getting-started/en/): Ini adalah panduan untuk mengatur dan menggunakan MySQL.
* [Memulai dengan Microsoft SQL Server](https://docs.microsoft.com/en-us/sql/relational-databases/tutorial-getting-started-with-the-database-engine?view=sql-server-ver15): Ini adalah tutorial untuk mulai menggunakan SQL Server.
* [Memulai PostgreSQL](https://www.postgresql.org/docs/10/tutorial-start.html): Ini adalah tutorial untuk mulai menggunakan PostgreSQL.
* [Memulai dengan SQLite](https://www.sqlite.org/quickstart.html): Ini adalah panduan memulai cepat untuk menggunakan SQLite.

**Panduan mendalam: praktik terbaik SQL**

*Anda dapat menyimpan bacaan ini untuk referensi di masa mendatang. Silakan mengunduh versi PDF dari bacaan di bawah ini:*

[Panduan mendalam DAC3\_ praktik terbaik SQL.pdf](https://d3c33hcgiwev3.cloudfront.net/5vVDkB5qT1y1Q5Aeau9c_Q_6d0e31160e2e43479d172390d19853f1_DAC3-In-depth-guide_-SQL-best-practices.pdf?Expires=1693872000&Signature=GcTwQsogkkghBiOTeL31CJFEINhtxg1p-IXUE~GmvQw3sv0R9IAlPJ1pbtydir37742cw8JQwa3x0O44xlA7a-EW-cl0o0u4lqJ2oAGZDMhvQKGKYhJ8YBlyhxVGSs6JjicVziFpPraUsNx2HOXm8alUtxalUiRo3aj9I~5-15E_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A" \t "_blank)

[File pdf](https://d3c33hcgiwev3.cloudfront.net/5vVDkB5qT1y1Q5Aeau9c_Q_6d0e31160e2e43479d172390d19853f1_DAC3-In-depth-guide_-SQL-best-practices.pdf?Expires=1693872000&Signature=GcTwQsogkkghBiOTeL31CJFEINhtxg1p-IXUE~GmvQw3sv0R9IAlPJ1pbtydir37742cw8JQwa3x0O44xlA7a-EW-cl0o0u4lqJ2oAGZDMhvQKGKYhJ8YBlyhxVGSs6JjicVziFpPraUsNx2HOXm8alUtxalUiRo3aj9I~5-15E_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A" \t "_blank)

Praktik terbaik ini mencakup pedoman untuk menulis kueri SQL, mengembangkan dokumentasi, dan contoh yang menunjukkan praktik ini. Ini adalah sumber yang bagus untuk berguna saat Anda menggunakan SQL sendiri; Anda bisa langsung membuka bagian yang relevan untuk meninjau praktik ini. Anggap saja seperti panduan lapangan SQL!

**Kapitalisasi dan sensitivitas huruf besar-kecil**

Dengan SQL, penggunaan huruf besar biasanya tidak menjadi masalah. Anda bisa menulis SELECT atau select atau SeLeCT. Semuanya berfungsi! Namun jika Anda menggunakan huruf besar sebagai bagian dari gaya yang konsisten, kueri Anda akan terlihat lebih profesional.

Untuk menulis kueri SQL seperti seorang profesional, sebaiknya gunakan huruf besar semua sebagai permulaan klausa (misalnya, SELECT, FROM, WHERE, dll.). Fungsi juga harus menggunakan huruf kapital semua (misalnya, SUM()). Nama kolom harus menggunakan huruf kecil semua (lihat bagian Snake\_case nanti dalam panduan ini). Nama tabel harus dalam CamelCase (lihat bagian CamelCase nanti dalam panduan ini). Hal ini membantu menjaga kueri Anda tetap konsisten dan lebih mudah dibaca tanpa memengaruhi data yang akan diambil saat Anda menjalankannya. Satu-satunya saat kapitalisasi menjadi penting adalah saat berada di dalam tanda kutip (selengkapnya tentang tanda kutip di bawah).

Vendor database SQL mungkin menggunakan variasi SQL yang sedikit berbeda. Variasi ini disebut **dialek SQL** . Beberapa dialek SQL peka terhadap huruf besar-kecil. BigQuery adalah salah satunya. Vertica adalah hal lain. Namun sebagian besar, seperti MySQL, PostgreSQL, dan SQL Server, tidak peka huruf besar-kecil. Ini berarti jika Anda mencari country\_code = 'us', semua entri yang memiliki 'us', 'us', 'Us', dan 'US' akan dikembalikan. Hal ini tidak terjadi pada BigQuery. BigQuery peka terhadap huruf besar-kecil, sehingga penelusuran yang sama hanya akan menghasilkan entri yang kode\_negaranya persis 'kami'. Jika kode\_negara adalah 'AS', BigQuery tidak akan mengembalikan entri tersebut sebagai bagian dari hasil Anda.

**Tanda kutip tunggal atau ganda: '' atau " "**

Pada umumnya, tidak masalah jika Anda menggunakan tanda kutip tunggal ' ' atau tanda kutip ganda " " saat mengacu pada string. Misalnya, SELECT adalah klausa starter. Jika Anda memasukkan SELECT dalam tanda kutip seperti 'SELECT' atau "SELECT", maka SQL akan memperlakukannya sebagai string teks. Kueri Anda akan menghasilkan kesalahan karena kueri Anda memerlukan klausa SELECT.

Namun ada dua situasi yang menentukan jenis kutipan yang Anda gunakan:

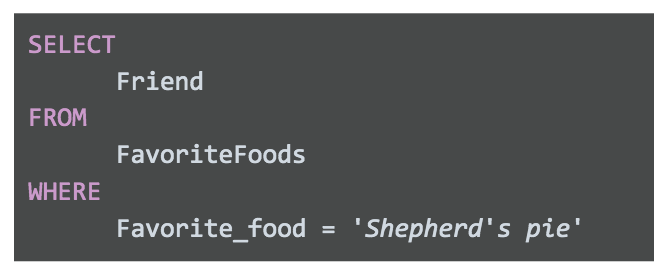
1. Saat Anda ingin string dapat diidentifikasi dalam dialek SQL  *apa pun*
2. Ketika string Anda berisi tanda kutip atau tanda kutip

Dalam setiap dialek SQL terdapat aturan tentang apa yang diterima dan apa yang tidak. Namun aturan umum di hampir semua dialek SQL adalah menggunakan tanda kutip tunggal untuk string. Ini membantu menghilangkan banyak kebingungan. Jadi jika kita ingin mereferensikan negara US dalam klausa WHERE (misalnya, kode\_negara = 'US'), maka gunakan tanda kutip tunggal di sekitar string 'US'.

Situasi kedua adalah ketika string Anda memiliki tanda kutip di dalamnya. Misalkan Anda memiliki kolom makanan favorit di tabel bernama Makanan Favorit dan kolom lainnya berhubungan dengan masing-masing teman.

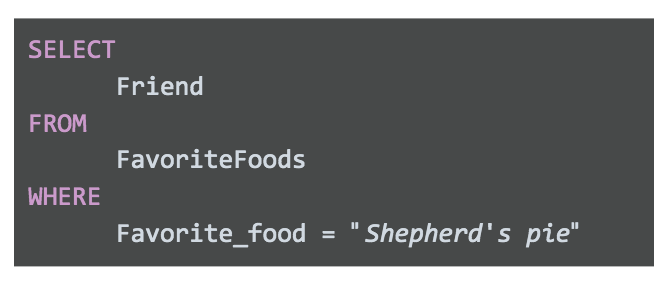
| **Teman** | **Makanan favorit** |
| --- | --- |
| Rachel DeSantos | kue gembala |
| Sujin Lee | Taco |
| Najil Okoro | paella Spanyol |

Anda mungkin memperhatikan makanan favorit Rachel mengandung tanda kutip. Jika Anda menggunakan tanda kutip tunggal dalam klausa WHERE untuk mencari teman yang memiliki makanan favorit tersebut, tampilannya akan seperti ini:



**Ini tidak akan berhasil.** Jika Anda menjalankan kueri ini, Anda akan mendapatkan pesan kesalahan sebagai balasannya. Hal ini karena SQL mengenali string teks sebagai sesuatu yang dimulai dengan tanda kutip 'dan diakhiri dengan tanda kutip lain'. Jadi dalam kueri buruk di atas, SQL menganggap Makanan\_Favorit yang Anda cari adalah 'Gembala'. Hanya 'Gembala' karena apostrof di Shepherd **mengakhiri** string.

Secara umum, ini adalah satu-satunya saat Anda menggunakan tanda kutip ganda, bukan tanda kutip tunggal. Jadi kueri Anda akan terlihat seperti ini:



SQL memahami string teks dimulai dengan tanda kutip tunggal ' atau tanda kutip ganda". Karena string ini dimulai dengan tanda kutip ganda, SQL akan mengharapkan tanda kutip ganda lainnya untuk menandakan akhir dari string. Ini menjaga tanda kutip tetap aman, sehingga akan mengembalikan " Kue Shepherd" dan bukan 'Gembala'.

**Komentar sebagai pengingat**

Saat Anda merasa lebih nyaman dengan SQL, Anda akan dapat membaca dan memahami pertanyaan secara sekilas. Namun tidak ada salahnya untuk memiliki komentar dalam kueri untuk mengingatkan diri Anda tentang apa yang Anda coba lakukan. Dan jika Anda membagikan pertanyaan Anda, itu juga membantu orang lain memahaminya.

Misalnya:

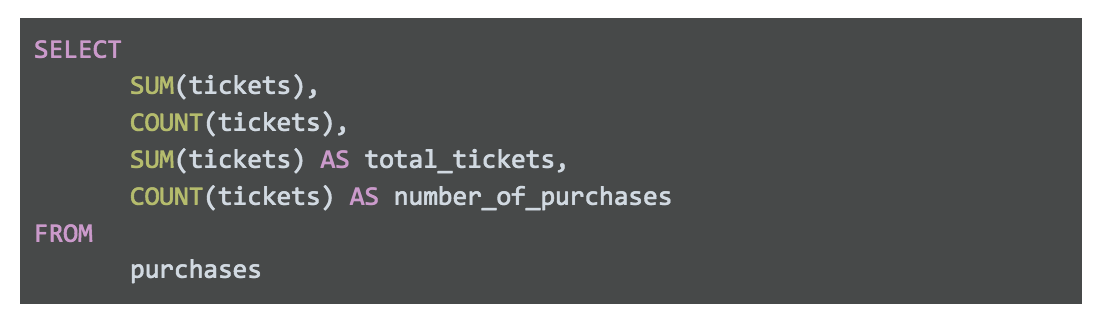
A screen shot of a computer code

Description automatically generated

Anda dapat menggunakan # sebagai pengganti dua tanda hubung, --, pada kueri di atas namun perlu diingat bahwa # tidak dikenali di semua dialek SQL (MySQL tidak mengenali #). Jadi yang terbaik adalah menggunakan -- dan konsisten dengannya. Saat Anda menambahkan komentar ke kueri menggunakan --, mesin kueri database akan mengabaikan semua yang ada di baris yang sama setelah --. Ini akan terus memproses kueri mulai dari baris berikutnya.

**Nama Snake\_case untuk kolom**

Penting untuk selalu memastikan bahwa keluaran kueri Anda memiliki nama yang mudah dipahami. Jika Anda membuat kolom baru (misalnya dari penghitungan atau penggabungan kolom baru), kolom baru tersebut akan menerima nama default umum (misalnya, f0). Misalnya:

Tabel berikut menampilkan hasil kueri ini: f0: 8 f1: 4 total\_tiket: 8 Jumlah\_pembelian\_: 4

Hasilnya adalah:

| **f0** | **f1** | **total\_tiket** | **jumlah\_pembelian\_** |
| --- | --- | --- | --- |
| 8 | 4 | 8 | 4 |

Dua kolom pertama diberi nama f0 dan f1 karena tidak disebutkan dalam kueri di atas. SQL defaultnya adalah f0, f1, f2, f3, dan seterusnya. Kami menamai dua kolom terakhir total\_tickets dan number\_of\_purchases sehingga nama kolom ini muncul di hasil kueri. Inilah sebabnya mengapa sebaiknya selalu memberi nama yang berguna pada kolom Anda, terutama saat menggunakan fungsi. Setelah menjalankan kueri, Anda ingin dapat dengan cepat memahami hasil Anda, seperti dua kolom terakhir yang kami jelaskan dalam contoh.

Selain itu, Anda mungkin memperhatikan bagaimana nama kolom memiliki garis bawah di antara kata-katanya. Nama tidak boleh mempunyai spasi di dalamnya. Jika 'total\_tickets' memiliki spasi dan tampak seperti 'total tiket' maka SQL akan mengganti nama SUM(tiket) menjadi 'total' saja. Karena spasinya, SQL akan menggunakan 'total' sebagai namanya dan tidak akan mengerti apa yang Anda maksud dengan 'tiket'. Jadi, spasi buruk dalam nama SQL. Jangan pernah menggunakan spasi.

Praktik terbaiknya adalah menggunakan Snake\_case. Artinya 'total tiket', yang memiliki spasi di antara kedua kata tersebut, harus ditulis sebagai 'total\_tickets' dengan garis bawah, bukan spasi.

**Nama CamelCase untuk tabel**

Anda juga dapat menggunakan kapitalisasi CamelCase saat memberi nama tabel Anda. Kapitalisasi CamelCase berarti Anda menggunakan huruf kapital di awal setiap kata, seperti unta berpunuk dua (Baktria). Jadi tabel TicketsByOccasion menggunakan kapitalisasi CamelCase. Harap diperhatikan bahwa penggunaan huruf besar pada kata pertama di CamelCase bersifat *opsional;* camelCase juga digunakan. Beberapa orang membedakan kedua gaya tersebut dengan memanggil CamelCase, PascalCase, dan memesan camelCase ketika kata pertama tidak menggunakan huruf kapital, seperti unta berpunuk satu (Dromedary); misalnya, ticketByOccasion.

Pada akhirnya, CamelCase adalah pilihan gaya. Ada cara lain untuk memberi nama tabel Anda, termasuk:

* Semua huruf kecil atau besar, seperti ticketbyoccasion atau TICKETSBYOCCASION
* Dengan ular\_kasus, seperti ticket\_by\_occasion

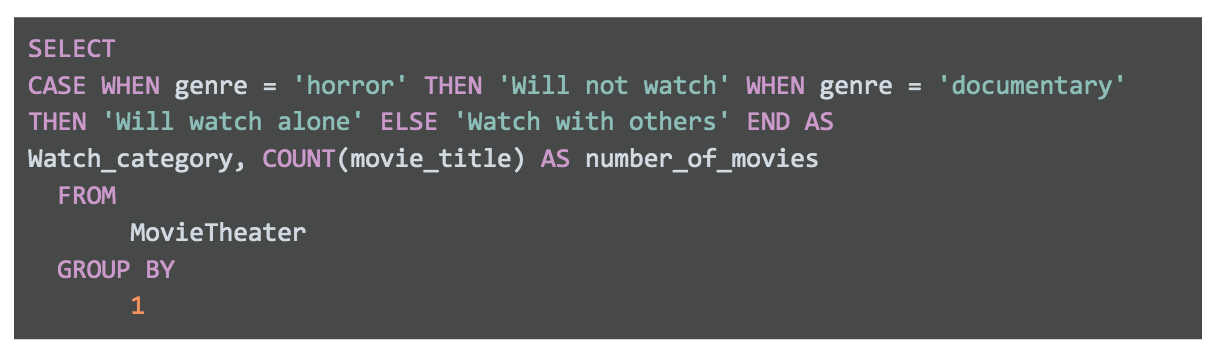
Perlu diingat, opsi dengan huruf kecil atau besar semua dapat mempersulit pembacaan nama tabel Anda, sehingga tidak disarankan untuk penggunaan profesional.

Opsi kedua, Snake\_case, secara teknis oke. Dengan kata-kata yang dipisahkan dengan garis bawah, nama tabel Anda mudah dibaca, namun bisa menjadi sangat panjang karena Anda menambahkan garis bawah. Menulis juga membutuhkan lebih banyak waktu. Jika Anda sering menggunakan meja ini, itu bisa menjadi tugas.

Singkatnya, terserah Anda untuk menggunakan Snake\_case atau CamelCase saat membuat nama tabel. Pastikan saja nama tabel Anda mudah dibaca dan konsisten. Pastikan juga untuk mengetahui apakah perusahaan Anda memiliki cara yang disukai untuk memberi nama tabel mereka. Jika ya, selalu gunakan konvensi penamaan mereka untuk konsistensi.

**Lekukan**

Sebagai aturan umum, Anda ingin menjaga panjang setiap baris dalam kueri <= 100 karakter. Ini membuat pertanyaan Anda mudah dibaca. Misalnya, periksa kueri ini dengan baris >100 karakter:

PILIH KASUS WHEN genre = 'horror' THEN 'Tidak akan menonton' WHEN genre = 'dokumenter' THEN 'Akan menonton sendiri' ELSE 'Tonton bersama orang lain' END AS Watch\_category, COUNT(

Kueri ini sulit dibaca dan sulit dipecahkan atau diedit. Sekarang, inilah kueri yang kita gunakan untuk tetap berpegang pada aturan <= 100 karakter:

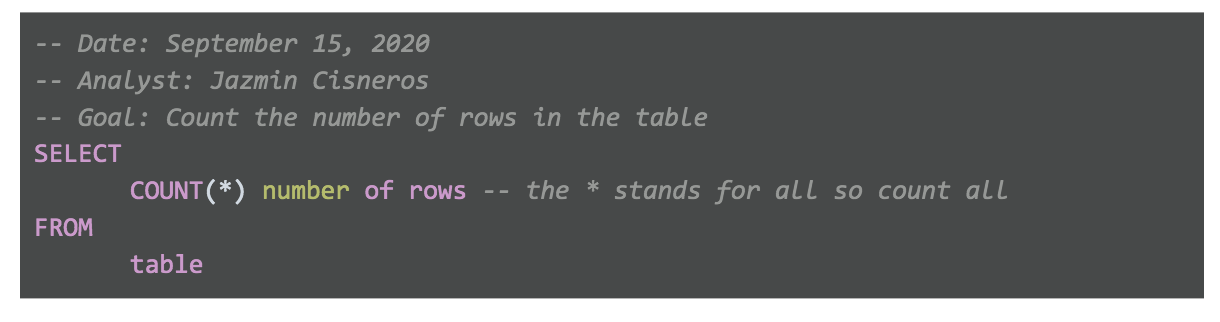
A screen shot of a computer code

Description automatically generated

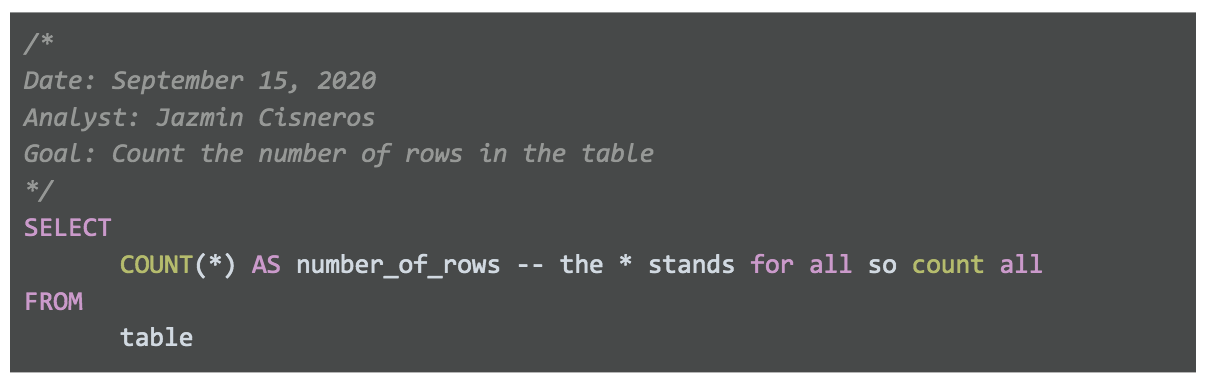
Sekarang lebih mudah untuk memahami apa yang Anda coba lakukan dalam klausa SELECT. Tentu, kedua kueri akan berjalan tanpa masalah karena lekukan tidak menjadi masalah dalam SQL. Namun lekukan yang tepat tetap penting untuk menjaga garis tetap pendek. Dan itu akan dihargai oleh siapa pun yang membaca pertanyaan Anda, termasuk Anda sendiri!

**Komentar multi-baris**

Jika Anda membuat komentar yang memerlukan beberapa baris, Anda dapat menggunakan -- untuk setiap baris. Atau, jika Anda memiliki lebih dari dua baris komentar, mungkin lebih rapi dan mudah menggunakan /\* untuk memulai komentar dan \*/ untuk menutup komentar. Misalnya, Anda dapat menggunakan metode -- seperti di bawah ini:

-- Tanggal: 15 September 2020 -- Analis: Jazmin Cisneros -- Sasaran: Menghitung jumlah baris dalam tabel SELECT COUNT(\*) jumlah baris -- tanda \* berarti semua jadi hitung semua FROM tabel

Atau, Anda dapat menggunakan metode /\* \*/ seperti di bawah ini:

/\* Tanggal: 15 September 2020 Analis: Jazmin Cisneros Sasaran: Menghitung jumlah baris dalam tabel \*/ SELECT COUNT(\*) AS number\_of\_rows -- \* berarti semua jadi hitung semua FROM tabel

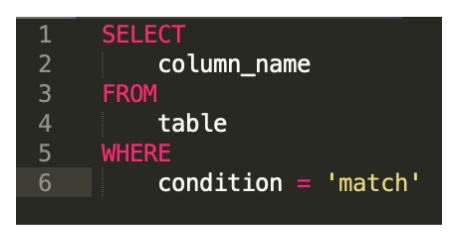
Dalam SQL, tidak masalah metode mana yang Anda gunakan. SQL mengabaikan komentar apa pun yang Anda gunakan: #, --, atau /\* dan \*/. Jadi terserah Anda dan preferensi pribadi Anda. Metode /\* dan \*/ untuk komentar multi-baris biasanya terlihat lebih rapi dan membantu memisahkan komentar dari kueri. Namun tidak ada satu metode yang benar atau salah.

**Editor teks SQL**

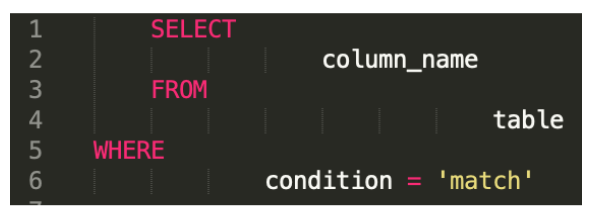
Saat Anda bergabung dengan sebuah perusahaan, Anda dapat mengharapkan setiap perusahaan menggunakan platform SQL dan dialek SQL mereka sendiri. Platform SQL yang mereka gunakan (misalnya BigQuery, MySQL, atau SQL Server) adalah tempat Anda menulis dan menjalankan kueri SQL Anda. Namun perlu diingat bahwa tidak semua platform SQL menyediakan editor skrip asli untuk menulis kode SQL. Editor teks SQL memberi Anda antarmuka tempat Anda dapat menulis kueri SQL dengan cara yang lebih mudah dan diberi kode warna. Faktanya, semua kode yang telah kita kerjakan sejauh ini ditulis dengan editor teks SQL!

**Contoh dengan Sublime Text**

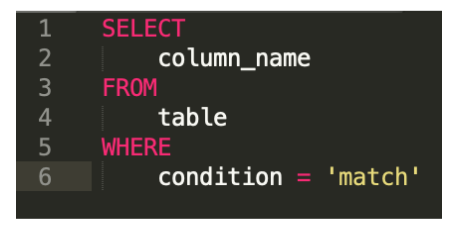
Jika platform SQL Anda tidak memiliki kode warna, Anda mungkin ingin mempertimbangkan untuk menggunakan editor teks seperti[Teks Luhur](https://www.sublimetext.com/)atau[Atom](https://atom.io/). Bagian ini menunjukkan bagaimana SQL ditampilkan di Sublime Text. Berikut ini query di Sublime Text:



Dengan Sublime Text, Anda juga dapat melakukan pengeditan lanjutan seperti menghapus indentasi di beberapa baris secara bersamaan. Misalnya, kueri Anda memiliki indentasi di tempat yang salah dan terlihat seperti ini:



Ini sangat sulit dibaca, jadi sebaiknya hilangkan indentasi tersebut dan mulai lagi dari awal. Dalam platform SQL biasa, Anda harus masuk ke setiap baris dan menekan BACKSPACE untuk menghapus setiap indentasi per baris. Namun di Sublime, Anda dapat menghilangkan semua indentasi secara bersamaan dengan memilih semua baris dan menekan Command (atau CTRL di Windows) + [. Ini menghilangkan indentasi dari setiap baris. Kemudian Anda dapat memilih baris yang ingin Anda indentasi (yaitu baris 2, 4, dan 6) dengan menekan tombol Command (atau tombol CTRL di Windows) dan memilih baris tersebut. Kemudian sambil tetap menahan tombol Command (atau tombol CTRL di Windows), tekan ] untuk membuat indentasi baris 2, 4, dan 6 secara bersamaan. Ini akan membersihkan kueri Anda dan membuatnya terlihat seperti ini:



Sublime Text juga mendukung ekspresi reguler. **Ekspresi reguler** (atau **regex** ) dapat digunakan untuk mencari dan mengganti pola string dalam kueri. Kami tidak akan membahas ekspresi reguler di sini, namun Anda mungkin ingin mempelajarinya sendiri lebih lanjut karena ini adalah alat yang sangat ampuh.

Anda dapat memulai dengan sumber daya berikut:

* [Cari dan ganti di Sublime Text](https://sublime-text-unofficial-documentation.readthedocs.io/en/latest/search_and_replace/search_and_replace_overview.html)
* [Tutorial ekspresi reguler](https://www.regular-expressions.info/tutorialcnt.html) (jika Anda tidak tahu apa itu ekspresi reguler)
* [Lembar contekan regex](https://jdhao.github.io/2019/02/28/sublime_text_regex_cheat_sheet/)