**Menggunakan SQL sebagai analis data junior**

Dalam bacaan ini, Anda akan mempelajari lebih lanjut tentang cara memutuskan kapan menggunakan SQL, atau Structured Query Language. Sebagai seorang analis data, Anda akan ditugaskan untuk menangani banyak data, dan SQL adalah salah satu alat yang dapat membantu mempermudah pekerjaan Anda. SQL adalah cara utama analis data mengekstrak data dari database. Sebagai seorang analis data, Anda akan bekerja dengan database sepanjang waktu, itulah sebabnya SQL adalah keterampilan utama. Mari ikuti bagaimana analis data junior menggunakan SQL untuk menyelesaikan tugas bisnis.

**Tugas dan konteks bisnis**

Analis data junior dalam contoh ini bekerja di perusahaan media sosial. Model bisnis baru diterapkan pada tanggal 15 Februari 2020 dan perusahaan ingin memahami pertumbuhan pengguna mereka dibandingkan tahun sebelumnya. Secara khusus, analis data diminta untuk mengetahui berapa banyak pengguna yang bergabung sejak 15 Februari 2020.



**Fungsi dan rumus spreadsheet atau kueri SQL?**

Sebelum mereka dapat menjawab pertanyaan ini, analis data ini perlu memilih alat apa yang akan digunakan. Pertama, mereka harus memikirkan di mana data tersebut berada. Jika disimpan dalam database, maka SQL adalah alat terbaik untuk pekerjaan itu. Namun jika disimpan dalam spreadsheet, maka mereka harus melakukan analisisnya di spreadsheet tersebut. Dalam skenario tersebut, mereka dapat membuat tabel pivot data lalu menerapkan rumus dan filter tertentu pada datanya hingga mereka mengetahui jumlah pengguna yang bergabung setelah tanggal 15 Februari. Ini bukanlah proses yang rumit, tetapi memerlukan banyak langkah.

Dalam hal ini, data disimpan dalam database, sehingga harus bekerja dengan SQL. Dan analis data ini mengetahui bahwa mereka bisa mendapatkan hasil yang sama dengan satu kueri SQL:

PILIH JUMLAH(BEDA user\_id) SEBAGAI count\_of\_unique\_users DARI tabel WHERE join\_date >= '2020-02-15'

Spreadsheet dan SQL keduanya memiliki kelebihan dan kekurangan:

| **Fitur Spreadsheet** | **Fitur Database SQL** |
| --- | --- |
| Kumpulan data yang lebih kecil | Kumpulan data yang lebih besar |
| Masukkan data secara manual | Akses tabel di seluruh database |
| Buat grafik dan visualisasi dalam program yang sama | Siapkan data untuk analisis lebih lanjut di perangkat lunak lain |
| Pemeriksaan ejaan bawaan dan fungsi berguna lainnya | Fungsionalitas yang cepat dan kuat |
| Paling baik saat bekerja sendiri dalam sebuah proyek | Cocok untuk kerja kolaboratif dan melacak kueri yang dijalankan oleh semua pengguna |

Dalam hal ini, lokasi data akan menentukan alat mana yang Anda gunakan. Jika Anda bekerja dengan data yang sudah ada dalam spreadsheet, kemungkinan besar di situlah Anda akan melakukan analisis. Dan jika Anda bekerja dengan data yang disimpan dalam database, SQL akan menjadi alat terbaik yang dapat Anda gunakan untuk analisis Anda. Anda akan mempelajari lebih lanjut tentang SQL yang akan datang, sehingga Anda siap menangani masalah bisnis apa pun dengan alat terbaik.

# Dialek SQL dan kegunaannya

Dalam bacaan ini, Anda akan mempelajari lebih lanjut tentang dialek SQL dan beberapa kegunaannya yang berbeda. Sebagai penyegaran singkat, **Structured Query Language** , atau SQL, adalah bahasa yang digunakan untuk berkomunikasi dengan database. Mempelajari SQL mirip seperti mempelajari bahasa baru — termasuk fakta bahwa bahasa biasanya memiliki dialek berbeda di dalamnya. Beberapa produk database memiliki varian SQL-nya sendiri, dan variasi dialek SQL yang berbeda inilah yang membantu Anda berkomunikasi dengan setiap produk database.

Dialek ini akan berbeda dari satu perusahaan ke perusahaan lain dan mungkin berubah seiring waktu jika perusahaan berpindah ke sistem database lain. Jadi, banyak analis memulai dengan SQL Standar dan kemudian menyesuaikan dialek yang mereka gunakan berdasarkan database yang mereka gunakan. SQL Standar berfungsi dengan sebagian besar database dan memerlukan sedikit perubahan sintaksis untuk beradaptasi dengan dialek lain.

Sebagai analis data junior, penting untuk mengetahui bahwa ada sedikit perbedaan antar dialek. Namun dengan menguasai SQL Standar, yaitu dialek yang akan Anda gunakan dalam program ini, Anda akan siap menggunakan SQL di database mana pun.

## Informasi lebih lanjut

Anda mungkin tidak perlu mengetahui setiap dialek SQL, namun ada gunanya mengetahui bahwa dialek-dialek yang berbeda ini ada. Jika Anda tertarik untuk mempelajari lebih lanjut tentang dialek SQL dan kapan dialek tersebut digunakan, Anda dapat memeriksa sumber daya berikut untuk informasi lebih lanjut:

* blog LearnSQL,[**Apa Itu Dialek SQL, dan Yang Mana yang Harus Anda Pelajari?**](https://learnsql.com/blog/what-sql-dialect-to-learn/)
* Artikel Bantuan Pengujian Perangkat Lunak,[**Perbedaan Antara SQL Vs MySQL vs SQL Server**](https://www.softwaretestinghelp.com/sql-vs-mysql-vs-sql-server/)
* Blog Kamp Data,[**SQL Server, PostgreSQL, MySQL... apa bedanya? Di mana saya memulai?**](https://www.datacamp.com/community/blog/sql-differences)Perhatikan bahwa ada kesalahan pada artikel blog ini. Tabel perbandingan salah menyatakan bahwa SQLite menggunakan subkueri alih-alih fungsi jendela. Mengacu kepada[**Fungsi Jendela SQLite**](https://sqlite.org/windowfunctions.html)dokumentasi untuk klarifikasi yang tepat.
* tutorial Tutorial SQL,[**Apa itu SQL**](https://www.sqltutorial.org/what-is-sql/)

**Optional: Upload the customer dataset to BigQuery**

In the next video, the instructor uses a specific dataset. The instructions in this reading are provided for you to upload the same dataset in your BigQuery console.

You must have a BigQuery account to follow along. If you have hopped around courses, [Using BigQuery](https://www.coursera.org/learn/data-preparation/supplement/DYOQK/using-bigquery) in the **Prepare Data for Exploration** course covers how to set up a BigQuery account.

**Prepare for the next video**

* First, download the CSV file from the attachment below.

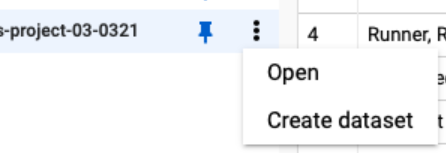
[Customer Table - Sheet1](https://d3c33hcgiwev3.cloudfront.net/F0iSyYcLT9iIksmHCw_Y-Q_191f150a80d74fda96f9df2aa2e3b533_Customer-Table---Sheet1.csv?Expires=1694044800&Signature=PEAQmf7uraejxBtFFCbE7Of-uT2QjQumfjwJ9uzI3ScPShuImUbRlxS0Yz7~X0D0OlZ5TTFB8fsKCt3hV47cpLfYNOEcJg-fE9sihM--C~zQ9MeHm67zljFOLf1GrdrhxC4f~lD1rWDvGsf6Vy6kYVsMMXHRQk5jXa4SosZ-tIk_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A" \t "_blank)

[CSV File](https://d3c33hcgiwev3.cloudfront.net/F0iSyYcLT9iIksmHCw_Y-Q_191f150a80d74fda96f9df2aa2e3b533_Customer-Table---Sheet1.csv?Expires=1694044800&Signature=PEAQmf7uraejxBtFFCbE7Of-uT2QjQumfjwJ9uzI3ScPShuImUbRlxS0Yz7~X0D0OlZ5TTFB8fsKCt3hV47cpLfYNOEcJg-fE9sihM--C~zQ9MeHm67zljFOLf1GrdrhxC4f~lD1rWDvGsf6Vy6kYVsMMXHRQk5jXa4SosZ-tIk_&Key-Pair-Id=APKAJLTNE6QMUY6HBC5A" \t "_blank)

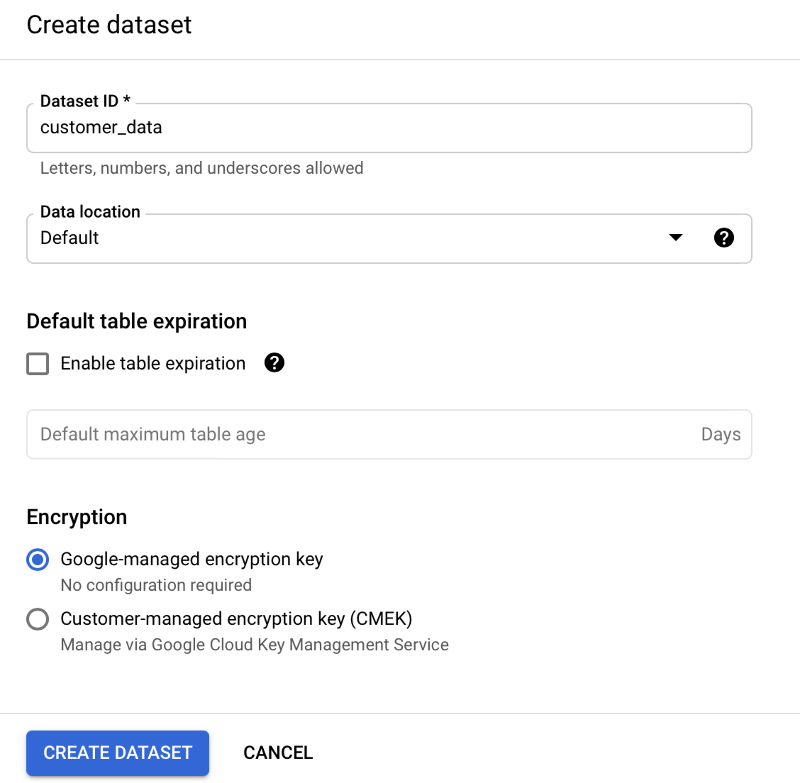
* Next, complete the following steps in your BigQuery console to upload the Customer Table dataset.

**Step 1**: Open your BigQuery console and click on the project you want to upload the data to.

**Step 2:** In the Explorer on the left, click the Actions icon (three vertical dots) next to your project name and select **Create dataset**.



**Step 3:** In the upcoming video, the name "customer\_data" will be used for the dataset. If you plan to follow along with the video, enter **customer\_data** for the Dataset ID.



**Step 4:** Click **CREATE DATASET** (blue button) to add the dataset to your project.

**Step 5:** In the Explorer on the left, click to expand your project, and then click the **customer\_data** dataset you just created.

**Step 6:** Click the Actions icon (three vertical dots) next to customer\_data and select **Open**.

**Step 7:** Click the blue **+** icon at the middle to open the Create table window.

Click the create table button.

**Step 8:** Under Source, for the Create table from selection, choose where the data will be coming from.

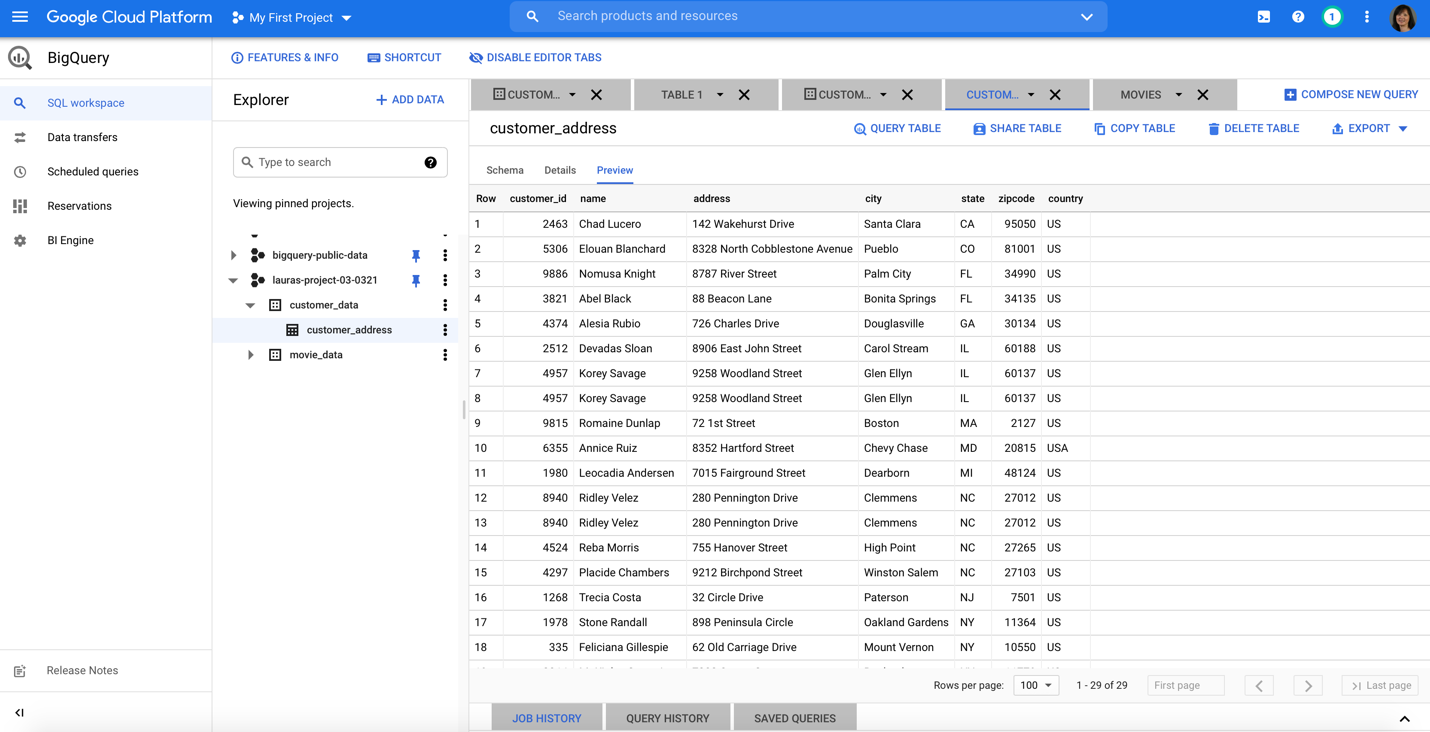
* Select **Upload**.
* Click **Browse** to select the Customer Table CSV file you downloaded.
* Choose **CSV** from the file format drop-down.

**Step 9:** For Table name, enter **customer\_address** if you plan to follow along with the video.

**Step 10:** For Schema, click the Auto detect check box.

**Step 11:** Click **Create** **table** (blue button). You will now see the **customer\_address** table under your **customer\_data** dataset in your project.

**Step 12:** Click **customer\_address** and then select the Preview tab. Confirm that you see the data shown below.



And now you have everything you need to follow along with the next video. This is also a great table to use to practice querying data on your own. Plus, you can use these steps to upload any other data you want to work with.

Pengalaman SQL

* Mendpatkan data dari table menggunakan pernyataan SELECT
* Hapus duplikat data menggunakan perintah DISTINCT dan COUNT + WHERE
* Memanipulasi data string dengan TRIM() dan SUBSTR.
* Membuat/menjatuhkan table dengan CREATE TABLE dan DROP TABLE
* Mengubah tipe data dengan CAST.