

Formatting for better analysis

Combine multiple datasets

- ✓

**Reading:** Import and combine data in spreadsheets and databases  
20 min
- ✓

**Reading:** Step-by-Step: Merge text strings to gain insights  
20 min
- ▶

**Video:** Merge text strings to gain insights  
4 min
- 📖

**Practice Quiz:** Hands-On Activity: Combine multiple pieces of data  
1h
- ✓

**Reading:** Step-by-Step: Strings in spreadsheets  
20 min
- ▶

**Video:** Strings in spreadsheets  
3 min
- ✓

**Reading:** Manipulate strings with SQL  
10 min
- 🔗

**Ungraded Plugin:** SQL query syntax  
10 min
- 📖

**Practice Quiz:** Test your knowledge on combining multiple datasets  
8 min

Get support during analysis

Module 2 challenge

# Manipulate strings with SQL

An important part of a data analyst's job is knowing how to convert and manipulate data for analysis. One way data analysts manipulate strings is to concatenate them, which means to join together two or more text strings. Once strings are concatenated, they form a new, longer text string for analysis. In this reading, you'll learn about different SQL functions that can be used to concatenate strings.

## CONCAT in action

Here are some examples of how you might use **CONCAT** as you work with data.

### CONCAT

You're working with the marketing team on an email campaign, and you need to generate full names from your database's first and last name columns. SQL's **CONCAT** function allows you to join together two or more string values, simplifying this task, as follows:

```
1  SELECT
2      CONCAT(first_name, ' ', last_name) AS full_name
3  FROM
4      customers;
```

In this example, **CONCAT** merges the **first\_name** and **last\_name** fields to create a new field called **full\_name**. The space ( ' ') separator ensures the full name appears properly.

### CONCAT\_WS

Now, you're tasked with creating a report that includes a website's URL components: the protocol (http), domain name (**your\_company**), and domain (**com**). You'd use **CONCAT\_WS**, which stands for **CONCAT With Separator**, to achieve this. It's similar to **CONCAT**, but it includes a separator, such as a space or period, between the strings.

```
1  SELECT CONCAT_WS('.', 'www', 'your_company', 'com') as website FROM web_data;
```

Here, **CONCAT\_WS** adds a period ( ' . ') between each part of the website URL, ensuring the URL is in the correct, navigable format.

### CONCAT with ||

In BigQuery, you can use the || operator to concatenate strings. For instance, if you're working with a dataset of book information and want to create a full title by combining the book's name and its edition, you could use ||, like so:

```
1  SELECT book_name || ' - ' || edition AS full_book_title FROM library;
```

This script combines the book name and edition, separated by a hyphen for clarity, providing a complete, informative title for your records.

**Note:** In some other SQL environments, you cannot use the || operator to concatenate strings. You must use + instead. For example, to concatenate the strings 'Google' and '.com' in Microsoft SQL server, you would use:

```
1  SELECT 'Google' + '.com'
```

Always ensure you're using the correct syntax for the specific SQL environment you're working in!

## Concatenate strings with SQL

Review the table below as a summary of the **CONCAT** function and its variations in SQL.

| Function/<br>operator | Use   | Example  | Result         |
|-----------------------|---|--|----------------|
| CONCAT                | Concatenate strings to create new text strings                                | CONCAT('Google',<br>' .com')                   | Google.com     |
| CONCAT_WS             | Concatenate two or more strings together with a separator between each string | CONCAT_WS(' . ',<br>'www', 'google',<br>'com') | www.google.com |
|                       | Concatenate two or more strings together with the    operator                 | 'Google'    '.com'                             | Google.com     |

## Key takeaways

In SQL, **CONCAT** is a function that joins strings together to create new text strings. This is useful for creating new variables or features for data analysis, as well as more readable and informative output. In this way, **CONCAT** can simplify your data analysis and make you more efficient.