**Get started with data modeling, schemas, and databases**

**Choose the right database**

**How data moves**

**Data-processing with Dataflow**

**Organize data in BigQuery**

**Video:** Gather information from stakeholders
3 min

**Reading:** Merge data from multiple sources with BigQuery
10 min

**Practice Quiz:** Activity: Set up a sandbox and query a public dataset in BigQuery
4 questions

**Reading:** Unify data with target tables
10 min

**Practice Quiz:** Activity: Create a target table in BigQuery
3 questions

**Reading:** Activity Exemplar: Create a target table in BigQuery
20 min

**Reading:** Case study: Wayfair - Working with stakeholders to create a pipeline
20 min

**Review: Data models and pipelines**

**[Optional] Review Google Data Analytics Certificate content**

# Merge data from multiple sources with BigQuery

Previously, you started exploring Google Dataflow, a Google Cloud Platform (GCP) tool that reads data from the source, transforms it, and writes it in the destination location. In this lesson, you will begin working with another GCP data-processing tool: BigQuery. As you may recall from the Google Data Analytics Certificate, BigQuery is a data warehouse used to query and filter large datasets, aggregate results, and perform complex operations.

As a business intelligence (BI) professional, you will need to gather and organize data from stakeholders across multiple teams. BigQuery allows you to merge data from multiple sources into a target table. The target table can then be turned into a dashboard, which makes the data easier for stakeholders to understand and analyze. In this reading, you will review a scenario in which a BI professional uses BigQuery to merge data from multiple stakeholders in order to answer important business questions.

## The problem

Consider a scenario in which a BI professional, Aviva, is working for a fictitious coffee shop chain. Each year, the cafes offer a variety of seasonal menu items. Company leaders are interested in identifying the most popular and profitable items on their seasonal menus so that they can make more confident decisions about pricing; strategic promotion; and retaining, expanding, or discontinuing menu items.

## The solution

### Data extraction

In order to obtain the information the stakeholders are interested in, Aviva begins extracting the data. The data extraction process includes locating and identifying relevant data, then preparing it to be transformed and loaded. To identify the necessary data, Aviva implements the following strategies:

#### Meet with key stakeholders

Aviva leads a workshop with stakeholders to identify their objectives. During this workshop, she asks stakeholders questions to learn about their needs:

- What information needs to be obtained from the data (for instance, performance of different menu items at different restaurant locations)?
- What specific metrics should be measured (sales metrics, marketing metrics, product performance metrics)?
- What sources of data should be used (sales numbers, customer feedback, point of sales)?
- Who needs access to this data (management, market analysts)?
- How will key stakeholders use this data (for example, to determine which items to include on upcoming menus, make pricing decisions)?

#### Observe teams in action

Aviva also spends time observing the stakeholders at work and asking them questions about what they're doing and why. This helps her connect the goals of the project with the organization's larger initiatives. During these observations, she asks questions about why certain information and activities are important for the organization.

### Organize data in BigQuery

Once Aviva has completed the data extraction process, she transforms the data she's gathered from different stakeholders and loads it into BigQuery. Then she uses BigQuery to design a target table to organize the data. The target table helps Aviva unify the data. She then uses the target table to develop a final dashboard for stakeholders to review.

## The results

When stakeholders review the dashboard, they are able to identify several key findings about the popularity and profitability of items on their seasonal menus. For example, the data indicates that many peppermint-based products on their menus have decreased in popularity over the past few years, while cinnamon-based products have increased in popularity. This finding leads stakeholders to decide to retire three of their peppermint-based drinks and bakery items. They also decide to add a selection of new cinnamon-based offerings and launch a campaign to promote these items.

## Key findings

Organizing data from multiple sources in a tool like BigQuery allows BI professionals to find answers to business questions. Consolidating the data in a target table also makes it easier to develop a dashboard for stakeholders to review. When stakeholders can access and understand the data, they can make more informed decisions about how to improve services or products and take advantage of new opportunities.