Inspect a dataset: A guided, hands-on tour

As a data analyst, you'll use data to answer questions and solve problems. When you analyze data and draw conclusions, you are generating insights that can influence business decisions, drive positive change, and help your stakeholders meet their goals.

Before you begin an analysis, it's important to inspect your data to determine if it contains the specific information you need to answer your stakeholders' questions. In any given dataset, it may be the case that:

- The data is not there (you have sandwich data, but you need pizza data)
- The data is insufficient (you have pizza data for June 1-7, but you need data for the entire month of June)
- The data is incorrect (your pizza data lists the cost of a slice as $250, which makes you question the validity of the dataset)

Inspecting your dataset will help you pinpoint what questions are answerable and what data is still missing. You may be able to recover this data from an external source or at least recommend to your stakeholders that another data source be used.

In this reading, imagine you're a data analyst inspecting spreadsheet data to determine if it's possible to answer your stakeholders' questions.

The scenario

You are a data analyst working for an ice cream company. Management is interested in improving the company's ice cream sales.

The company has been collecting data about its sales—but not a lot. The available data is from an internal data source and is based on sales for 2019. You've been asked to review the data and provide some insight into the company's ice cream sales. Ideally, management would like answers to the following questions:

1. What is the most popular flavor of ice cream?
2. How does temperature affect sales?
3. How do weekends and holidays affect sales?
4. How does profitability differ for new versus returning customers?