

STUDY GUIDE

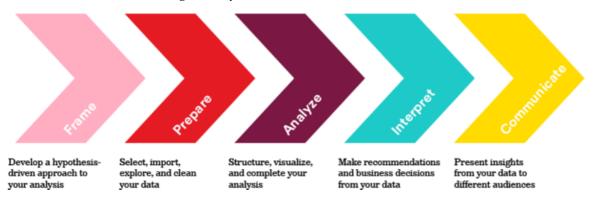
## **WHAT IS DATA SCIENCE?**

## **Key Terms and Definitions**

**Data Science:** An interdisciplinary, problem-solving-oriented framework comprising a set of tools and techniques used to extract useful information from data, as well as an application of scientific techniques to practical problems.

The Data Science Workflow: The process used to work through and complete data science projects, made up of the following steps:

- Frame: Develop a hypothesis-driven approach to your analysis.
- Prepare: Select, import, explore, and clean your data.
- Analyze: Structure, visualize, and complete your analysis.
- Interpret: Create recommendations and business decisions from your data.
- Communicate: Present insights from your data to different audiences.



• Data Dictionary: Describes the types and content of a data set.

## **Guiding Questions**

- 1. Why is the data science workflow said to be "nonlinear?"
- 2. Why is it helpful for a data scientist to have subject matter expertise?
- 3. Give an example of a framing question for a data science project relevant to your work.