

## STUDY GUIDE

# FINDING THE RIGHT DATA

---

## Data Sources

Data can be collected in several different ways:

- **First-party data:** Internal data that the client has collected directly.
- **Second-party data:** Data from a partner who has directly collected it.
  - **Scraping:** A data science technique where Python or another program is used to search through the code of a website and pull data based on a set of criteria.
  - **Open-source data:** Data freely available to everyone to use. Kaggle is a great resource for data analysts with free open-source data sets and posted models and results.
- **Third-party data:** Data aggregated from multiple sources and sold by a data supplier for a fee.