

Data analysis is the practice of working with data to glean useful information, which can then be used to make informed decisions. Companies are wisening up to the benefits of leveraging data. Data analysis can help a bank to personalize customer interactions, a health care system to predict future health needs, or an entertainment company to create the next big streaming hit.

Data analysis process

As the data available to companies continues to grow both in amount and complexity, so too does the need for an effective and efficient process by which to harness the value of that data. The analysis method typically moves through several iterative phases. Let's take a closer look at each.

- Identify the business question you'd like to answer. What problem is the company trying to solve? What do you need to measure, and how will you measure it?
- Collect the raw data sets you'll need to help you answer the identified question. Data collection might come from internal sources, like a company's client relationship management (CRM) software, or from secondary sources, like government records or social media application programming interfaces (APIs).
- Clean the data to prepare it for analysis. This often involves purging duplicate and anomalous data, reconciling inconsistencies, standardizing data structure and format, and dealing with white spaces and other syntax errors.
- Analyze the data. By manipulating the data using various data analysis tools and techniques, you can begin to find trends, correlations, outliers, and variations that begin to tell a story. During this stage, you might use data mining to discover patterns within databases or data visualization software to help transform data into an easy-to-understand graphical format.
- Interpret the results of your analysis to see how well the data answered your original question. What recommendations can you make based on the data? What are the limitations to your conclusions?