Data Mining

Data mining is the process of analyzing large amounts of [data](https://techterms.com/definition/data) in order to discover patterns and other information. It is typically performed on [databases](https://techterms.com/definition/database), which store data in a structured format. By "mining" large amounts of data, hidden information can be discovered and used for other purposes.

Data Mining Examples

A credit card company might use data mining to learn more about their members' buying habits. By analyzing purchases from cardholders across the United States, the company may discover shopping habits for different demographics, such as age, race, and location. This information could be useful in offering individuals specific promotions. The same data may also reveal shopping patterns in different regions of the country. This information could be valuable to companies looking to advertise or start businesses in specific states.

Online services, such as [Google](https://techterms.com/definition/google) and [Facebook](https://techterms.com/definition/facebook), mine enormous amounts of data to provide targeted content and advertisements to their users. Google, for example, might analyze search [queries](https://techterms.com/definition/query) to discover popular searches for certain areas and move those to the top of the autocomplete list (the suggestions that appear as you type). By mining user activity data, Facebook might discover popular topics among different age groups and provide targeted ads based on this information.

While data mining is commonly used for marketing purposes, it has many other uses as well. For instance, healthcare companies may use data mining to discover links between certain genes and diseases. Weather companies can mine data to discover weather patterns that may help predict future meteorologic events. Traffic management institutions can mine automotive data to forecast future traffic levels and create appropriate plans for highways and streets.

Data Mining Requirements

Data mining requires two things — lots of data and lots of computing power. The more organized the data, the easier it is to mine it for useful information. Therefore it is important for any organization that wants to engage in data mining to be proactive in selecting what data to log and how to store it. When it comes to mining the data, [supercomputers](https://techterms.com/definition/supercomputer) and [computing clusters](https://techterms.com/definition/cluster) may be used to process [petabytes](https://techterms.com/definition/petabyte) of data.