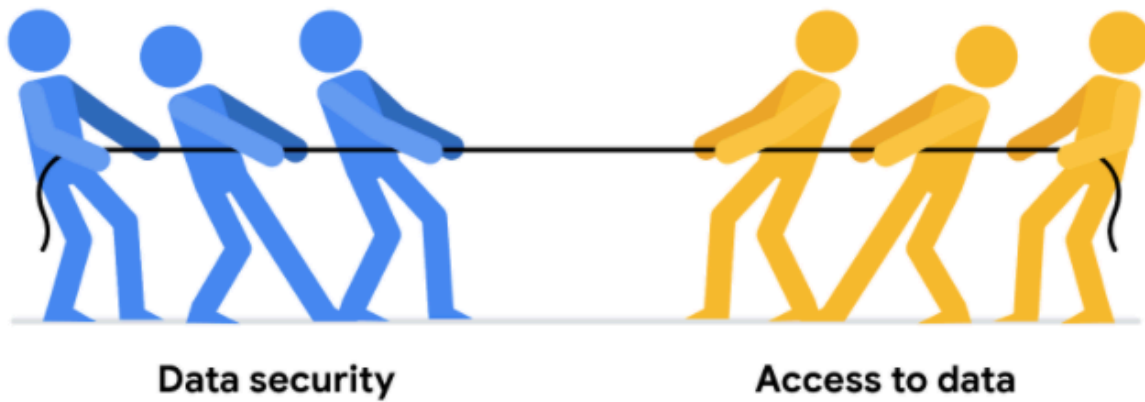


Balance security and analytics

Data security means protecting data from unauthorized access or corruption by putting safety measures in place. Usually the purpose of data security is to keep unauthorized users from accessing or viewing sensitive data. Data analysts have to find a way to balance data security with their actual analysis needs. This can be tricky-- we want to keep our data safe and secure, but we also want to use it as soon as possible so that we can make meaningful and timely observations. In order to do this, companies need to find ways to balance their data security measures with their data access needs.



Luckily, there are a few security measures that can help companies do just that. The two we will talk about here are encryption and tokenization.

Encryption uses a unique algorithm to alter data and make it unusable by users and applications that don't know the algorithm. This algorithm is saved as a "key" which can be used to reverse the encryption; so if you have the key, you can still use the data in its original form.

Tokenization replaces the data elements you want to protect with randomly generated data referred to as a "token." The original data is stored in a separate location and mapped to the tokens. To access the complete original data, the user or application needs to have permission to use the tokenized data and the token mapping. This means that even if the tokenized data is hacked, the original data is still safe and secure in a separate location.

Encryption and tokenization are just some of the data security options out there. There are a lot of others, like using authentication devices for AI technology.

As a junior data analyst, you probably won't be responsible for building out these systems. A lot of companies have entire teams dedicated to data security or hire third party companies that specialize in data security to create these systems. But it is important to know that all companies have a responsibility to keep their data secure, and to understand some of the potential systems your future employer might use.

However, one thing you absolutely can do to help strike the right balance is to use **version control** best practices. Version control enables all collaborators within a file to track changes over time. You can understand who made what changes to a file, when they were made, and why.

Here's a simple example: Perhaps you're working on a project with a team of other people. You are all collaborating within the same set of files, but each person is responsible for a different part of the project. Without version control, it would be very difficult to keep track of who made what changes to the files and when. This would lead to confusion and, even worse, people accidentally overwriting each other's work! Version control is essential for data analytics professionals because it allows users to effectively collaborate with others and experiment with new ideas without fear of losing their work.