GLOSSARY

Hadoop for Data Science Tips, Tricks, and Techniques

With Ben Sullins

Use these terms and definitions below to understand concepts taught in the course.

Transcript Search: note that you can search for terms directly within the course. To search video text, switch to the *Transcripts* tab, then press Cmd/Ctrl + F on your keyboard to run a search within the active transcript.

Term	Definition
aggregate table	A rollup table that summarizes the data for the user more quickly than if they were to query and analyze the data
flat file	A file type that usually has summary rows, header rows, and footer or tail rows on the bottom
HDFS	Hadoop Distributed File System; a data storage system that offers high-performance access to data
Impala	A faster query engine offered by Cloudera; useful when working with big data
partition	A tool used to separate data into specified sections
terminal window	A window in which users can type and execute text-based commands