Apache Parquet:

* Apache Parquet is a columnar storage format available to any project in the Hadoop ecosystem, regardless of the choice of data processing framework, data model or programming language.
* A column-oriented DBMS (or columnar database) is a database management system (DBMS) that stores data tables as columns rather than as rows.
* By storing data in columns rather than rows, the database can more precisely access the data it needs to answer a query rather than scanning and discarding unwanted data in rows.