

R FOR DATA SCIENCE

R FOR BIG DATA

SPARK



- Apache Spark is a **cluster computing platform** designed to be fast and general purpose.
- Spark extends the MapReduce model to support more types of computations, including **interactive queries** and **stream processing**.
- One of the main features Spark offers for speed is the ability to run computations **in memory**,
- The system is also more **efficient** than MapReduce for complex applications running on disk.