

Formatos de datos y casos de uso

- Soporta muchos formatos de datos:



CSV,TSV
(legibles por humanos)



JSON (legible
por humanos)



ORC
(columnar,
dividible)



Parquet
(columnar,
dividible)



Avro
(divisible)



Zlib, LZO, Gzip

- Casos de uso:



Análisis de logs



ETL (Extracción,
Transformación y
Carga)



Lagos de datos