

## Formatos de arquivo no contexto de big data

What is parquet? <https://databricks.com/glossary/what-is-parquet>

Should you use parquet? <https://blog.matthewrathbone.com/2019/12/20/parquet-or-bust.html>

Big Data File Formats <https://blog.clairvoyantsoft.com/big-data-file-formats-3fb659903271>

TOP 5 popular big data file formats: <https://youtu.be/fC-0Ubwo0gY>

Parquet vs AVRO vs ORC: <https://youtu.be/q8uqpwxavZw>

## Sistemas de arquivos

A closer look at Azure Data Lake Storage Gen2:

<https://azure.microsoft.com/en-in/blog/a-closer-look-at-azure-data-lake-storage-gen2/>

Storing Apache Hadoop Data on the Cloud - HDFS vs S3:

<https://www.xplenty.com/blog/storing-apache-hadoop-data-cloud-hdfs-vs-s3/>

File storage, blob storage, or object storage?

<https://www.redhat.com/en/topics/data-storage/file-block-object-storage>

A map of storage options in Google Cloud:

<https://cloud.google.com/blog/topics/developers-practitioners/map-storage-options-google-cloud>

## Spark

How Apache Spark's Transformations and Action Works:

[https://medium.com/@aristo\\_alex/how-apache-sparks-transformations-and-action-works-ceb0d03b00d0](https://medium.com/@aristo_alex/how-apache-sparks-transformations-and-action-works-ceb0d03b00d0)

Lazy Evaluation in Apache Spark – A Quick guide:

<https://data-flair.training/blogs/apache-spark-lazy-evaluation/>