Procesamiento de Datos Masivos

Material complementario Módulo 1

- Sitio oficial de GCP. https://cloud.google.com/?hl=es.

- Sitio oficial de AWS. https://aws.amazon.com/.

- Sitio oficial de Digital Ocean: https://www.digitalocean.com/products.

- What Are Transactional Databases, GCP. https://cloud.google.com/learn/what-are-transactional-databases.

- Sitio oficial de BigQuery. https://cloud.google.com/bigquery/docs/introduction.

- Evaluación de consultas en BigQuery. https://cloud.google.com/blog/products/bigquery/in-memory-query-execution-in-google-bigquery.

- Paper "**MapReduce: Simplified Data Processing on Large Clusters**". https://static.googleusercontent.com/media/research.google.com/en//archive/mapreduce-osdi04.pdf.

- Introduction to Apache Spark (AWS). https://aws.amazon.com/big-data/what-is-spark/.

- Learning Spark: Lightning-Fast Data Analytics (2nd Edition, O'Reilly). Damji, Wenig, Das & Lee.

- Documentación del core de PySpark: https://spark.apache.org/docs/latest/api/python/reference/pyspark.html#rdd-apis.

- Learning Spark: Lightning-Fast Data Analytics (2nd Edition, O'Reilly). Damji, Wenig, Das & Lee.