



---

**Data Science Indy**

---



## Commercial Spark Service

Founded by creators of Spark

Currently only supports AWS integration

(your own AWS account)

Trial - 30 day

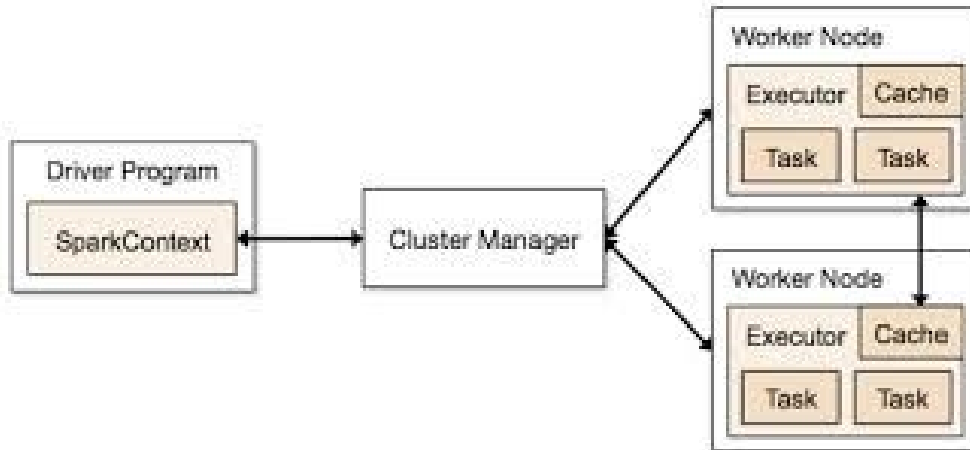
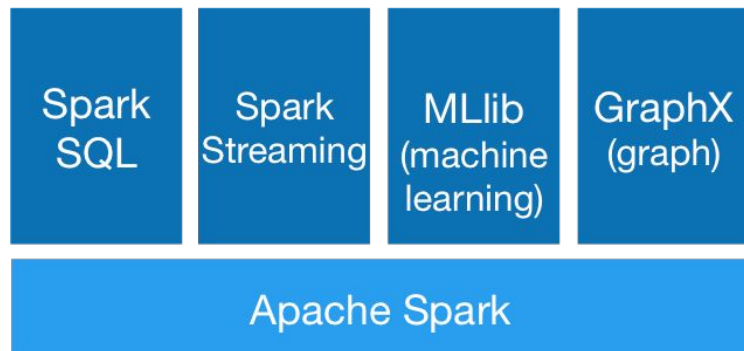
then \$249/mo (60GB Ram) +

\$0.35/hr (>30GB)

---

# Apache Spark Architecture

---



# databricks Demo

---

Cluster

Tables

Import files, database, S3

CSV, JSON, Avro, Parquet

**FORMAT COLUMN NAMES FOR HIVE BEFORE  
(no spaces)**

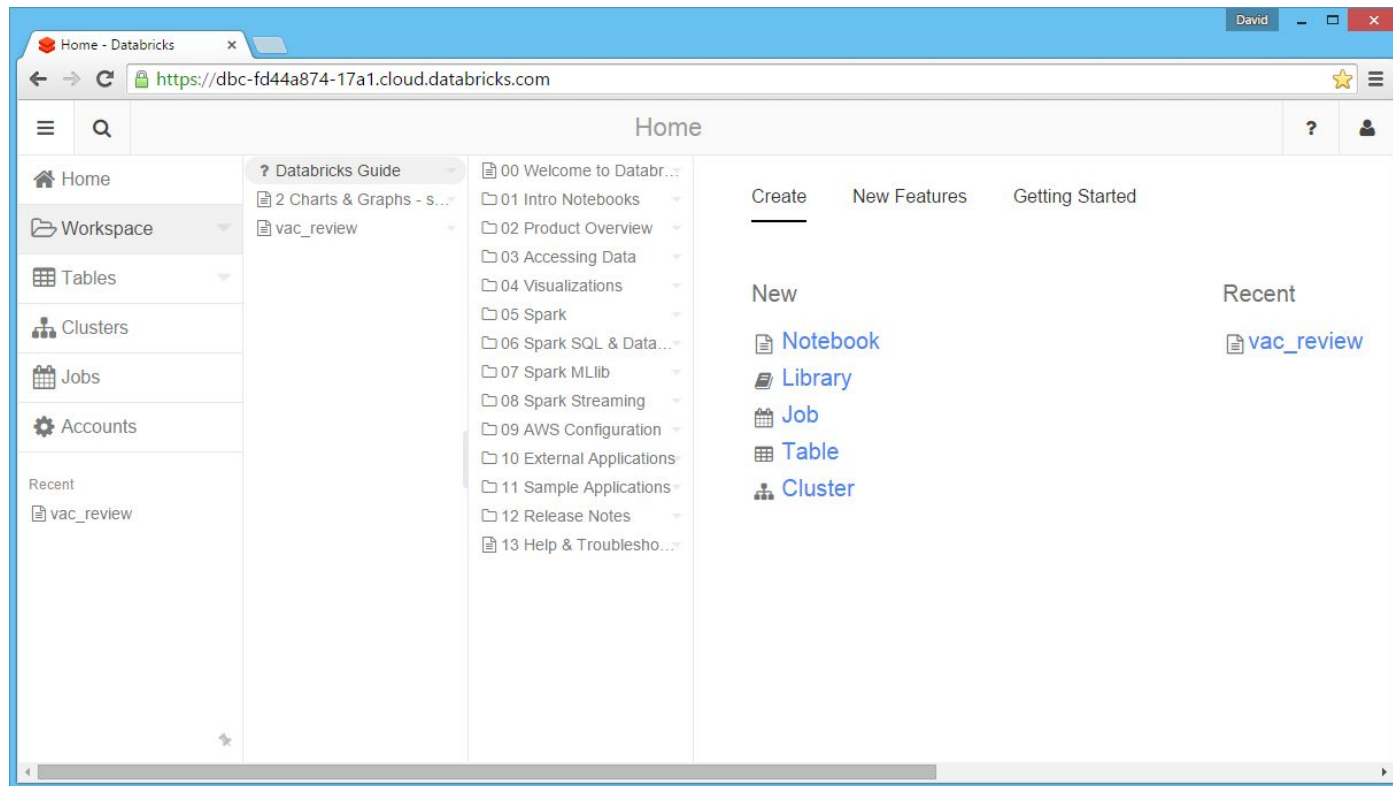
**KEEP COL TYPES STRING or DEFAULT**

Workspace

Notebooks

---

# databricks Demo cont.



# **databricks Demo cont.**

---

Workspace includes Demos

Tables

Clusters

Jobs

View Spark UI and Logs

---