SPARK Interface API

RDD	DataFrame	Dataset
Fault Tolerant	Fault Tolerant	Fault Tolerant
Distributed	Distributed	Distributed
Immutabilty	Immutabilty	Immutabilty
No schema	Schema	Schema
Slow on Non-JVM languages	Faster	Faster
No Execution optimization	optimization Catalyst optimizer	optimization
Low Level	High Level	High Level
No SQL Support	SQL Support	SQL Support
Type Safe	No type Safe	Type Safe
Syntax Error detected at Compile Time	Syntax Error detected at Compile Time	Syntax Error detected at Compile Time
Analysis Error Detected at Compile time	Analysis Error Detected at Run time	Analysis Error Detected at Compile time
JAVA,SCALA, Python,R	JAVA,SCALA, Python,R	JAVA, SCALA
Higher memory is used	Higher memory is used	Low memory is used. Tungsten encoders provide great benefits

We can seamlessly move between DataFrame or Dataset and RDDs by using simple syntax