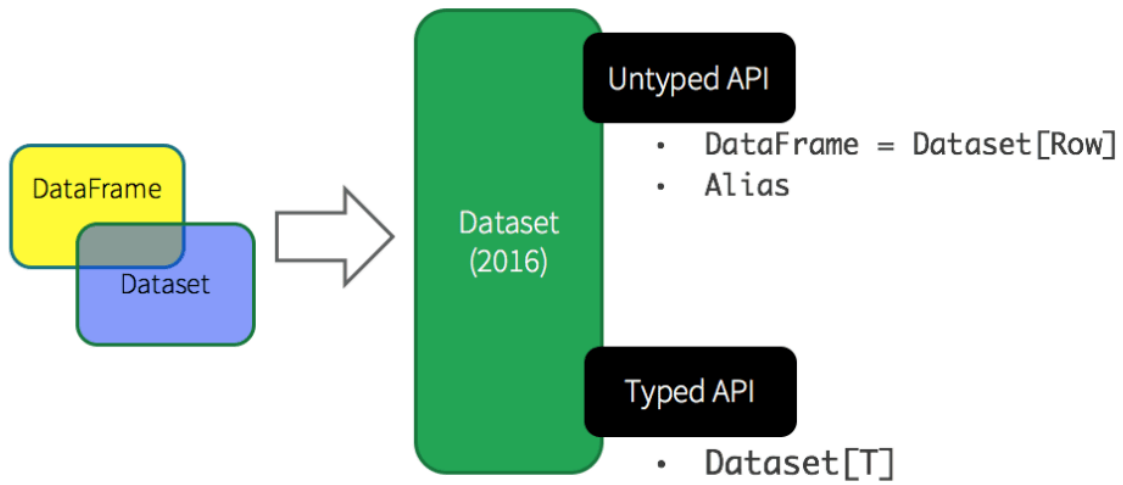Unified Apache Spark 2.0 API

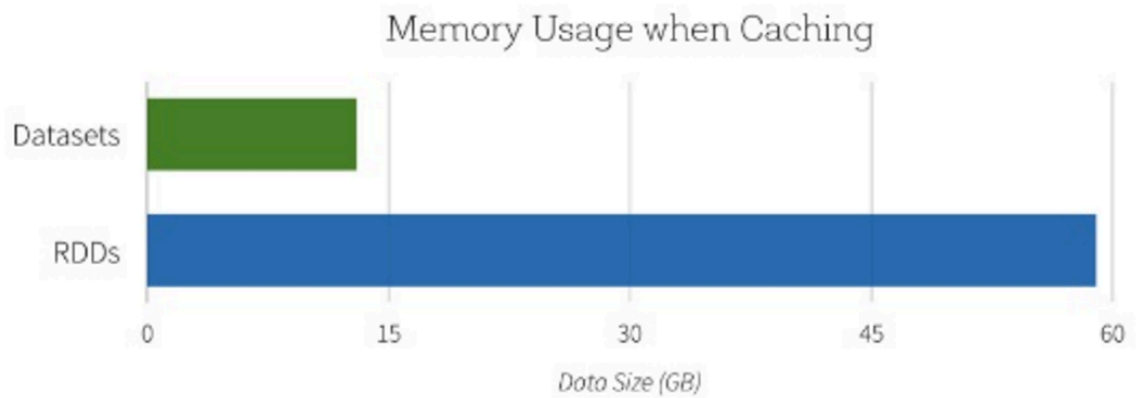## DataFrame

```
row.getAs("age")
```

DataSet

```
row.age
```

Space Efficiency

Memory Usage when Caching

Datasets	█████████████
RDDs	██

Data Size (GB)

0   15   30   45   60

```scala
import org.apache.spark.sql.{Encoders, SaveMode, SparkSession}

import java.util.Properties

private case class Person2(name:String, age:Long)

object SparkDsDf {
 def main(args: Array[String]): Unit = {

    val spark = SparkSession.builder.master("local").
      appName("SparkByExample")
      .getOrCreate()
    import spark.implicits._

    val personDf =
spark.read.json("/Users/serkan/Desktop/Course/Spark/data/peopleage.json")
    personDf.printSchema()
    personDf.show()

    personDf.foreach(row =>{
      println("Age : " + row.getLong(0) + ", Name : " + row.getString(1))
    })
    println("----------------------")
    personDf.foreach(row => {
      println("Age : " + row.getAs("age") + ", Name : " + row.getAs("name"))
    })

    val personDs = personDf.as[Person2]
    personDs.show()
    personDs.printSchema()

    personDs.foreach(row => {
      println("DS Age : " + row.age + ", Name : " + row.name)
    })
 }
}
```