```
scala> import org.apache.spark.ml.classification.LogisticRegression
import org.apache.spark.ml.classification.LogisticRegression

scala> import org.apache.spark.sql.SparkSession
import org.apache.spark.sql.SparkSession

scala>

scala> import org.apache.log4j._
import org.apache.log4j._

scala> Logger.getLogger("org").setLevel(Level.ERROR)
```

```
scala> data.printSchema()
root
 |-- Daily Time Spent on Site: double (nullable = true)
 |-- Age: integer (nullable = true)
 |-- Area Income: double (nullable = true)
 |-- Daily Internet Usage: double (nullable = true)
 |-- Ad Topic Line: string (nullable = true)
 |-- City: string (nullable = true)
 |-- Male: integer (nullable = true)
 |-- Country: string (nullable = true)
 |-- Timestamp: timestamp (nullable = true)
 |-- Clicked on Ad: integer (nullable = true)
```

```
scala> data.head(1)
res2: Array[org.apache.spark.sql.Row] = Array([68.95,35,61833.9,256.09,Cloned
5thgeneration orchestration,Wrightburgh,0,Tunisia,2016-03-27 00:53:11.0,0])

scala> val colnames = data.columns
colnames: Array[String] = Array(Daily Time Spent on Site, Age, Area Income, Daily
Internet Usage, Ad Topic Line, City, Male, Country, Timestamp, Clicked on Ad)

scala> val firstrow = data.head(1)(0)
firstrow: org.apache.spark.sql.Row = [68.95,35,61833.9,256.09,Cloned 5thgeneration
orchestration,Wrightburgh,0,Tunisia,2016-03-27 00:53:11.0,0]
```