
Question

[Niyaz Khafizov](#) · Sep 20, 2018

How to save spark dataframe via Spark connector

Hi all.

I want to insert my dataframe into InterSystems IRIS. So, I tried to do this:

```
df = spark.read.load("/home/imported-openssh-key/zeppelin-0.8.0-bin-all/bin/resultData3/DF.json",
format="json")
df.write.format("com.intersystems.spark")./
option("url", "IRIS://localhost:51773/DEDUPL")./
option("user", "*****").option("password", "*****")./
option("dbtable", "try.test1").save()
```

And got this error:

```
Py4JJavaError: An error occurred while calling o80.save.
: java.lang.NoSuchMethodError: org.apache.spark.sql.execution.datasources.jdbc.JdbcUtils$.schemaString(Lorg/apache/spark/sql/types/StructType;Ljava/lang/String;)Ljava/lang/String;
    at com.intersystems.spark.core.SparkAPI$class.schemaString(SparkAPI-2.1.scala:44)
    at com.intersystems.spark.core.DataSource.create$(DataSource.scala:237)
    at com.intersystems.spark.core.DataSource.createRelation(DataSource.scala:344)
    at org.apache.spark.sql.execution.datasources.SaveIntoDataSourceCommand.run(SaveIntoDataSourceCommand.scala:45)
    at org.apache.spark.sql.execution.command.ExecutedCommandExec.sideEffectResult$lzycompute(commands.scala:70)
    at org.apache.spark.sql.execution.command.ExecutedCommandExec.sideEffectResult(commands.scala:68)
    at org.apache.spark.sql.execution.command.ExecutedCommandExec.doExecute(commands.scala:86)
    at org.apache.spark.sql.execution.SparkPlan$$anonfun$execute$1.apply(SparkPlan.scala:131)
    at org.apache.spark.sql.execution.SparkPlan$$anonfun$execute$1.apply(SparkPlan.scala:127)
    at org.apache.spark.sql.execution.SparkPlan$$anonfun$executeQuery$1.apply(SparkPlan.scala:155)
    at org.apache.spark.rdd.RDDOperationScope$.withScope(RDDOperationScope.scala:151)
    at org.apache.spark.sql.execution.SparkPlan.executeQuery(SparkPlan.scala:152)
    at org.apache.spark.sql.execution.SparkPlan.execute(SparkPlan.scala:127)
    at org.apache.spark.sql.execution.QueryExecution.toRdd$lzycompute(QueryExecution.scala:80)
    at org.apache.spark.sql.execution.QueryExecution.toRdd(QueryExecution.scala:80)
    at org.apache.spark.sql.DataFrameWriter$$anonfun$runCommand$1.apply(DataFrameWriter.scala:654)
    at org.apache.spark.sql.DataFrameWriter$$anonfun$runCommand$1.apply(DataFrameWriter.scala:654)
    at org.apache.spark.sql.execution.SQLExecution$.withNewExecutionId(SQLExecution.scala:77)
```

```

at org.apache.spark.sql.DataFrameWriter.runCommand(DataFrameWriter.scala:654)
at org.apache.spark.sql.DataFrameWriter.saveToV1Source(DataFrameWriter.scala:273)
at org.apache.spark.sql.DataFrameWriter.save(DataFrameWriter.scala:267)
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
at java.lang.reflect.Method.invoke(Method.java:498)
at py4j.reflection.MethodInvoker.invoke(MethodInvoker.java:244)
at py4j.reflection.ReflectionEngine.invoke(ReflectionEngine.java:357)
at py4j.Gateway.invoke(Gateway.java:282)
at py4j.commands.AbstractCommand.invokeMethod(AbstractCommand.java:132)
at py4j.commands.CallCommand.execute(CallCommand.java:79)
at py4j.GatewayConnection.run(GatewayConnection.java:238)
at java.lang.Thread.run(Thread.java:748)

```

I tried this:

```

df = spark.read.load("/home/imported-openssh-key/zeppelin-0.8.0-bin-all/bin/resultData3/DF.json",
format="json")
df.write.format("iris")./
option("url", "IRIS://localhost:51773/DEDUPL")./
option("user", "*****").option("password", "*****")./
option("dbtable", "try.test1").save()

```

And got this error:

```

Py4JJavaError: An error occurred while calling o166.save.
: java.lang.NoSuchMethodError: org.apache.spark.sql.execution.datasources.jdbc.JdbcUtils$.schemaString(Lorg/apache/spark/sql/types/StructType;Ljava/lang/String;)Ljava/lang/String;
at com.intersystems.spark.core.SparkAPI$class.schemaString(SparkAPI-2.1.scala:44)
at com.intersystems.spark.core.DataSource.create$(DataSource.scala:237)
at com.intersystems.spark.core.DataSource.createRelation(DataSource.scala:344)
at org.apache.spark.sql.execution.datasources.SaveIntoDataSourceCommand.run(SaveIntoDataSourceCommand.scala:45)
at org.apache.spark.sql.execution.command.ExecutedCommandExec.sideEffectResult$lzycompute(commands.scala:70)
at org.apache.spark.sql.execution.command.ExecutedCommandExec.sideEffectResult(commands.scala:68)
at org.apache.spark.sql.execution.command.ExecutedCommandExec.doExecute(commands.scala:86)
at org.apache.spark.sql.execution.SparkPlan$$anonfun$execute$1.apply(SparkPlan.scala:131)
at org.apache.spark.sql.execution.SparkPlan$$anonfun$execute$1.apply(SparkPlan.scala:127)
at org.apache.spark.sql.execution.SparkPlan$$anonfun$executeQuery$1.apply(SparkPlan.scala:155)
at org.apache.spark.rdd.RDDOperationScope$.withScope(RDDOperationScope.scala:151)
at org.apache.spark.sql.execution.SparkPlan.executeQuery(SparkPlan.scala:152)
at org.apache.spark.sql.execution.SparkPlan.execute(SparkPlan.scala:127)
at org.apache.spark.sql.execution.QueryExecution.toRdd$lzycompute(QueryExecution.scala:80)
at org.apache.spark.sql.execution.QueryExecution.toRdd(QueryExecution.scala:80)
at org.apache.spark.sql.DataFrameWriter$$anonfun$runCommand$1.apply(DataFrameWriter.scala:654)

```

```

at org.apache.spark.sql.DataFrameWriter$$anonfun$runCommand$1.apply(DataFrameWriter.
scala:654)
at org.apache.spark.sql.execution.SQLExecution$.withNewExecutionId(SQLExecution.sca
a:77)
at org.apache.spark.sql.DataFrameWriter.runCommand(DataFrameWriter.scala:654)
at org.apache.spark.sql.DataFrameWriter.saveToV1Source(DataFrameWriter.scala:273)
at org.apache.spark.sql.DataFrameWriter.save(DataFrameWriter.scala:267)
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java
:43)
at java.lang.reflect.Method.invoke(Method.java:498)
at py4j.reflection.MethodInvoker.invoke(MethodInvoker.java:244)
at py4j.reflection.ReflectionEngine.invoke(ReflectionEngine.java:357)
at py4j.Gateway.invoke(Gateway.java:282)
at py4j.commands.AbstractCommand.invokeMethod(AbstractCommand.java:132)
at py4j.commands.CallCommand.execute(CallCommand.java:79)
at py4j.GatewayConnection.run(GatewayConnection.java:238)
at java.lang.Thread.run(Thread.java:748)

```

Also, I tried this:

```

df = spark.read.load("/home/imported-openssh-key/zeppelin-0.8.0-bin-all/bin/resultData3/DF.json",
format="json")
df.write.format("iris")./
option("driver","com.intersystems.jdbc.IRISDriver")./
option("url","IRIS://localhost:51773/DEDUPL")./
option("user","*****").option("password","*****")./
option("dbtable","try.test1").save()

```

And got this error:

```

Py4JJavaError: An error occurred while calling o210.save.
: java.lang.NoSuchMethodError: org.apache.spark.sql.execution.datasources.jdbc.JdbcUt
ils$.schemaString(Lorg/apache/spark/sql/types/StructType;Ljava/lang/String;)Ljava/lan
g/String;
at com.intersystems.spark.core.SparkAPI$class.schemaString(SparkAPI-2.1.scala:44)
at com.intersystems.spark.core.DataSource.create$1(DataSource.scala:237)
at com.intersystems.spark.core.DataSource.createRelation(DataSource.scala:344)
at org.apache.spark.sql.execution.datasources.SaveIntoDataSourceCommand.run(SaveInto
DataSourceCommand.scala:45)
at org.apache.spark.sql.execution.command.ExecutedCommandExec.sideEffectResult$lzyco
mpute(commands.scala:70)
at org.apache.spark.sql.execution.command.ExecutedCommandExec.sideEffectResult(comma
nds.scala:68)
at org.apache.spark.sql.execution.command.ExecutedCommandExec.doExecute(commands.sca
la:86)
at org.apache.spark.sql.execution.SparkPlan$$anonfun$execute$1.apply(SparkPlan.sca
la:131)
at org.apache.spark.sql.execution.SparkPlan$$anonfun$execute$1.apply(SparkPlan.sca
la:127)
at org.apache.spark.sql.execution.SparkPlan$$anonfun$executeQuery$1.apply(SparkPlan.
scala:155)
at org.apache.spark.rdd.RDDOperationScope$.withScope(RDDOperationScope.scala:151)
at org.apache.spark.sql.execution.SparkPlan.executeQuery(SparkPlan.scala:152)
at org.apache.spark.sql.execution.SparkPlan.execute(SparkPlan.scala:127)

```

```

at org.apache.spark.sql.execution.QueryExecution.toRdd$lzycompute(QueryExecution.scala:80)
at org.apache.spark.sql.execution.QueryExecution.toRdd(QueryExecution.scala:80)
at org.apache.spark.sql.DataFrameWriter$$anonfun$runCommand$1.apply(DataFrameWriter.scala:654)
at org.apache.spark.sql.DataFrameWriter$$anonfun$runCommand$1.apply(DataFrameWriter.scala:654)
at org.apache.spark.sql.execution.SQLExecution$.withNewExecutionId(SQLExecution.scala:77)
at org.apache.spark.sql.DataFrameWriter.runCommand(DataFrameWriter.scala:654)
at org.apache.spark.sql.DataFrameWriter.saveToV1Source(DataFrameWriter.scala:273)
at org.apache.spark.sql.DataFrameWriter.save(DataFrameWriter.scala:267)
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
at java.lang.reflect.Method.invoke(Method.java:498)
at py4j.reflection.MethodInvoker.invoke(MethodInvoker.java:244)
at py4j.reflection.ReflectionEngine.invoke(ReflectionEngine.java:357)
at py4j.Gateway.invoke(Gateway.java:282)
at py4j.commands.AbstractCommand.invokeMethod(AbstractCommand.java:132)
at py4j.commands.CallCommand.execute(CallCommand.java:79)
at py4j.GatewayConnection.run(GatewayConnection.java:238)
at java.lang.Thread.run(Thread.java:748)

```

And tried this:

```

df = spark.read.load("/home/imported-openssh-key/zeppelin-0.8.0-bin-all/bin/resultData3/DF.json",
format="json")
#format("iris").
df.write.format("com.intersystems.jdbc.IRISDriver")./
option("url", "IRIS://localhost:51773/DEDUPL")./
option("user", "*****").option("password", "*****")./
option("dbtable", "try.test1").save()

```

Got this error:

Py4JJavaError: An error occurred while calling o305.save.

```

: java.lang.RuntimeException: com.intersystems.jdbc.IRISDriver does not allow create
table as select.
at scala.sys.package$.error(package.scala:27)
at org.apache.spark.sql.execution.datasources.DataSource.planForWriting(DataSource.s
cala:550)
at org.apache.spark.sql.DataFrameWriter.saveToV1Source(DataFrameWriter.scala:278)
at org.apache.spark.sql.DataFrameWriter.save(DataFrameWriter.scala:267)
at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
at sun.reflect.NativeMethodAccessorImpl.invoke(NativeMethodAccessorImpl.java:62)
at sun.reflect.DelegatingMethodAccessorImpl.invoke(DelegatingMethodAccessorImpl.java:43)
at java.lang.reflect.Method.invoke(Method.java:498)
at py4j.reflection.MethodInvoker.invoke(MethodInvoker.java:244)
at py4j.reflection.ReflectionEngine.invoke(ReflectionEngine.java:357)
at py4j.Gateway.invoke(Gateway.java:282)
at py4j.commands.AbstractCommand.invokeMethod(AbstractCommand.java:132)
at py4j.commands.CallCommand.execute(CallCommand.java:79)
at py4j.GatewayConnection.run(GatewayConnection.java:238)

```

```
at java.lang.Thread.run(Thread.java:748)
```

So, how the one can save a dataframe to InterSystems IRIS?

[#AI](#) [#API](#) [#Machine Learning](#) [#InterSystems IRIS](#)

Source URL: <https://community.intersystems.com/post/how-save-spark-dataframe-spark-connector>