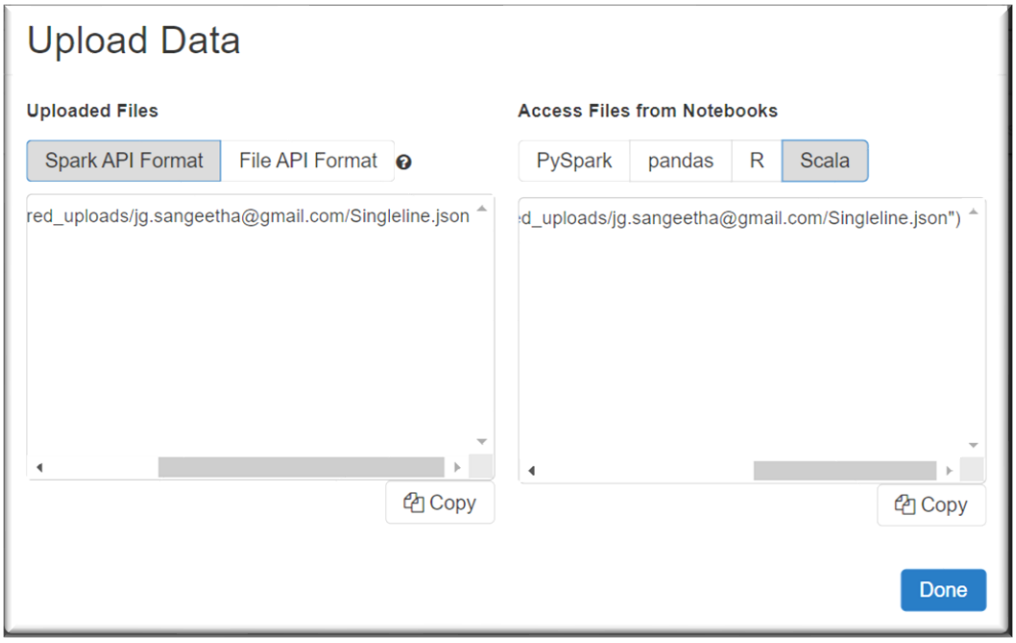
Spark Activity-3

Uploading json file using spark SQL data frames

1)Create Scala notebook and upload Singleline.json data. Copy the path of the uploaded data set.



2)Read the json file

3)Printing the schema

```
1  //Printing the schema
2  df.printSchema()

root
 |-- color: string (nullable = true)
 |-- value: string (nullable = true)
```

4) Save the file as temporary table

```
1  //saving as temporary table
2  df.createTempView("Json")
```

5)Select all data from temporary table and store it in a different variable

```
1  val data=sqlContext.sql("select * from Json")

▸ data:  org.apache.spark.sql.DataFrame = [color: string, value: string]
data: org.apache.spark.sql.DataFrame = [color: string, value: string]
```

6)Display all the records

```
1  data.show()

▸ (1) Spark Jobs
+-------+-----+
|  color|value|
+-------+-----+
|    red| #f00|
|  green| #0f0|
|   blue| #00f|
|   cyan| #0ff|
|magenta| #f0f|
| yellow| #ff0|
|  black| #000|
+-------+-----+
```

7)Follow step 1 to step 6 for processing multiline json file.