

Homework3_Simonsen

May 22, 2024

1 Assignment 3

1.1 Steven Simonsen

1.2 5/27/24

1.2.1 Instructions

Once you have successfully installed spark on your local machine you will need to execute the following code. Please keep in mind there may be libraries you will need to install which are included below. Once you have executed the code successfully please knit the document to a PDF and turn it in on Canvas.

SS Note - I used DataBricks for this lab. Entering this note as instructed to show evidence for DataBricks.

```
[0]: from datetime import datetime, date
from pyspark.sql import Row

df = spark.createDataFrame([
    Row(a=1, b=2., c='string1', d=date(2000, 1, 1), e=datetime(2000, 1, 1, 12, 0, 0)),
    Row(a=2, b=3., c='string2', d=date(2000, 2, 1), e=datetime(2000, 1, 2, 12, 0, 0)),
    Row(a=4, b=5., c='string3', d=date(2000, 3, 1), e=datetime(2000, 1, 3, 12, 0, 0))
])
df.show()
```

```
+---+---+-----+-----+-----+
|  a|  b|      c|      d|      e|
+---+---+-----+-----+-----+
|  1|2.0|string1|2000-01-01|2000-01-01 12:00:00|
|  2|3.0|string2|2000-02-01|2000-01-02 12:00:00|
|  4|5.0|string3|2000-03-01|2000-01-03 12:00:00|
+---+---+-----+-----+-----+
```

```
[0]:
```