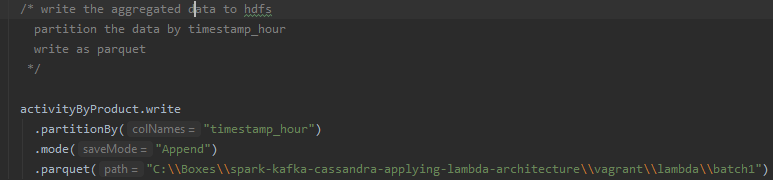
**WRITE THE AGGREGATED OUTPUT TO HDF LOCATION**

1.Write the aggregated data to a location under lambda

* The file is written in parquet format
* It is partitioned by timestamp\_hour



There was error in windows 10 for writing directly to hdfs .

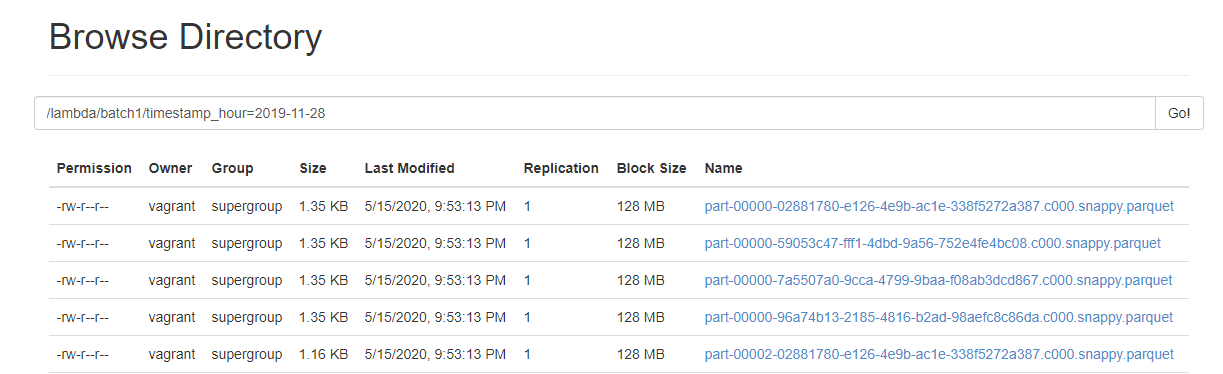
We need to

* Copy to the vagrant directory under windows as shown above
* ssh to the device and move the file to the warehouse

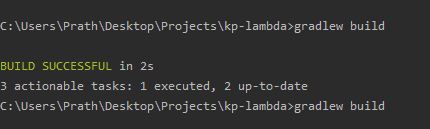
hdfs dfs -copyFromLocal /vagrant/lambda/batch1/ /lambda/

2. When we browse

<http://lambda-pluralsight:50070/explorer.html#/lambda/batch1/timestamp_hour=2019-11-28> under lambda we can see batch1 directory that was created for us and under that we can see that the data is portioned by time stamp



**3. build the project**



**3. We would find the jar under bin as shown below**

.bin/spark-submit --master yarn --deploy-mode cluster --class batch.BatchJob /vagrant/bigdata-kp-lambda.1.0-SNAPSHOT-shaded.jar

Note:This job failed as the development version and deployment version was diffent

**4.Vizualize the data in zeppelin**

