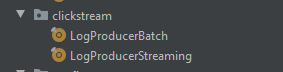
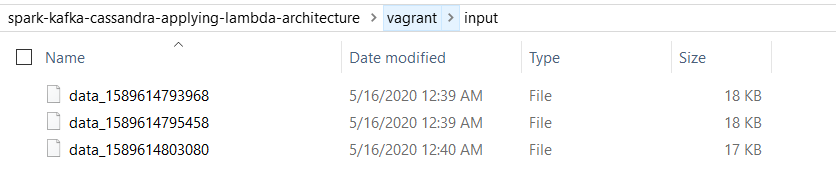
**GENERATE DATA FOR STREAMING**

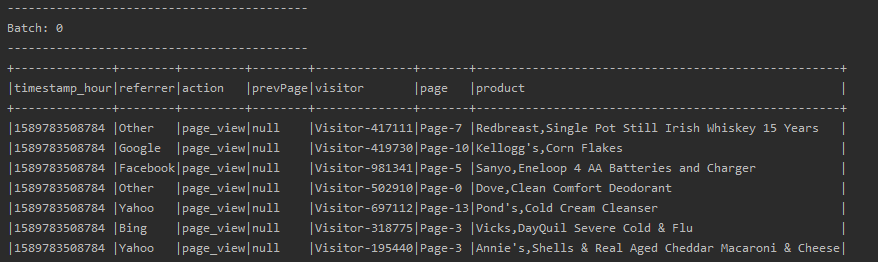
1. Destination file path and number of files has been specified

dest\_path = "C:\\Boxes\\spark-kafka-cassandra-applying-lambda-architecture\\vagrant\\input\\"  
number\_of\_files= 50

1. Add common.io for file operations
2. Change the log producer code to generate stream of files with delay
3. Run the log producer code and data would be generated with timestamp as shown below.

5. Create a new schema for activityStreams, read the stream and write to a console.

6. Run the StreamingJob



7. convert the data streams into a key value pairs

Key : combination and product and timestamp hour

Values: product details