I ran Kafka and Zookeeper on Linux Ubuntu over a Windows 10 machine. After installation of Java, Kafka and Zookeeper I performed the following steps.

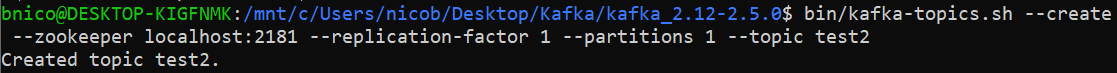
1. Start Zookeeper server in a first shell:



2. Open a second shell to start Kafka:



3. Open a third shell to create a topic, that consumers will use to receive messages from producers:



Start a consumer service (within the same 3rd shell):



4. Open a fourth shell to start a producer process:



5. I then successfully tested the producer to consumer messages both typing strings in the shell and inputting a simple .txt file. Next step included sending stream.jsonl to the producer. In the producer shell I executed the following command:



The result was a very long streaming of incomprehensible messages that after several minutes produced an error at the Kafka server.

6. The next thing to would be to find a way to parse the JSON file to understand the content, but I wasn’t successful in the due time. As for the data structure that I would use to count, probably a dictionary, considering that JSON format is based on keys.