

## Assignment 5 - Spark Streaming

1. Write a standalone application in Spark to count the number of words in a received stream.
  - **Hint 1:** use the *netcat* server to send the text through a TCP connection. For example, `nc -lk 9999` can be used to transfer any text that you type in the terminal through port 9999.
  - **Hint 2:** if you are using `local` as the master URL when creating `StreamingContext`, you have to give at least one more core as the number of input streams, because each input stream would create a receiver that occupies one core. If you use `local`, it only gives one core to the context, which is used by the socket stream's receiver, leaving no core available for processing the data, thus, you should use `local[2]`.
2. Extend the code to generate word count over last 30 seconds of data, and repeat the computation every 10 seconds.
3. Maintain a continuously updated word count for all the words in the stream.