HOW TO STREAM LOG DATA USING SPARK

Pre-Requisites of Flume Project:
hadoop-2.6.0
flume-1.6.0
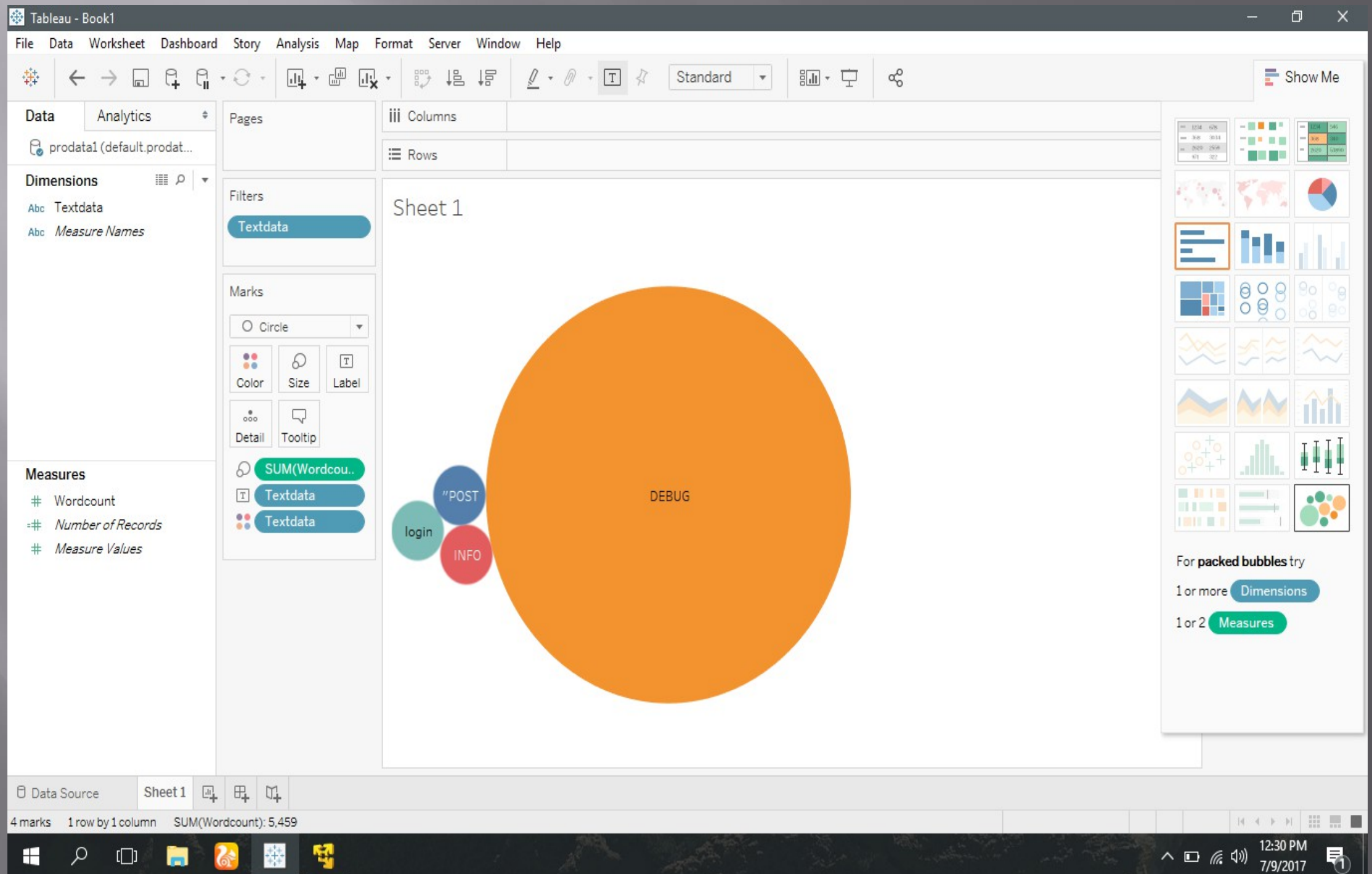
NOTE: Make sure that
install all the above components

▸**LEARNINGS OF THIS PROJECT:**

▸We will learn Flume Configurations and Commands

▸➢ Flume Agent

▸1. Source (Twitter Source)

▸2. Channel (Memory Channel)

▸3. Sink

▸Major project in Real Time `ServerLog Analysis

▸ 1. We are extracting the data from server logs 2.

▸This data will be useful to do analysis on Server Performance

▸Analyse the data using spark-scala ,Hive

▸Display the result in Data Visualization Tools

RESULTS
Final output

REFERENCES

- http://spark.apache.org/docs/latest/streaming-flume-integration.html/
- http://datametica.com/integration-of-spark-streaming-with-flume/
- https://www.sigmoid.com/getting-data-into-spark-streaming/