TA NOTES:

Troubleshooting Tips:

1. Make sure that the environment is proper i.e. the installation of jdk; and python/scala/java is done before running the code.
2. Always operate using ubuntu user by default. Use root permissions only when needed(e.g. to modify /etc/hosts, /etc/mysql/my.cnf).
3. Empty the buffers cache before any of your experiment run. You can run: sudo sh -c "sync; echo 3 > /proc/sys/vm/drop\_caches" on each slave VM.
4. **Debugging the Error**: Unable to convert Spark RDD to Schema RDD
   1. Example code written in object Temp: people.registerTempTable("people")
   2. No TypeTag available for Person at the line people.registerTempTable("people")
   3. Resolved by **Jira**: https://issues.apache.org/jira/browse/SPARK-4842