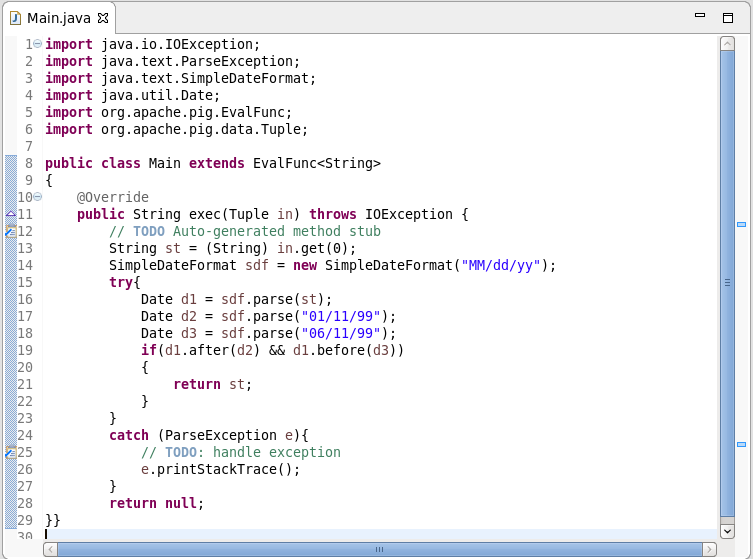
**Session 11**

**Assignment 6**

Find the top five kinds of Google Knowledge Occupation people who were guests in the show, in a particular time period i.e., between '1/11/99' and 6/11/99.

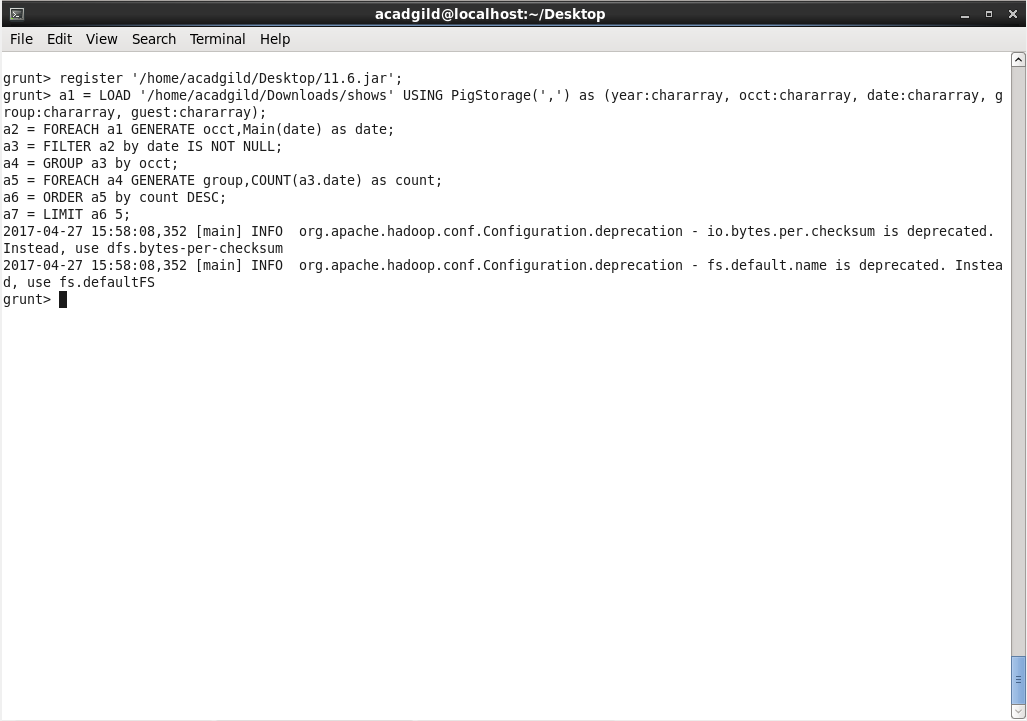
Need to write an UDF to compare the dates.

**UDF Program:**



To run this UDF program, the jar file is created and this jar file need to be registered in the Pig grunt shell.

**Pig Command lines:**



**Logic:**

**LINE 1:** The jar file ’11.6.jar’ is been registered to Pig shell using ‘register’ command.

**LINE 2:** Loading the dataset to pig grunt shell using ‘LOAD’ command.

**LINE 3:** Generating the column 2(Google Knowledge Occupation) and date after processing the dataset with UDF program.

**LINE 4:** Filtering the dataset by checking for ‘date’ is ‘not null’.

**LINE 5:** Grouping the dataset by column 2(Google Knowledge Occupation).

**LINE 6:** Generating the count for the dataset using date.

**LINE 7:** Ordering the dataset by the count.

**LINE 8:** Limiting the dataset up to first 5 entities.

**Output:**

