**athlete\_events.csv Dataset**

using Apache pig, we will try to gain more insights from these datasets.

Open pig in local mode as our all files are on the desktop and not in hdfs.

#pig –x local

## Problem Statement 1

Indians who have got gold medals till now along with sports name.

**Source Code**

|  |  |
| --- | --- |
|  | REGISTER 'piggybank.jar';  A = load ' athlete\_events.csv' USING org.apache.pig.piggybank.storage.CSVExcelStorage(',','NO\_MULTILINE','UNIX','SKIP\_INPUT\_HEADER');  B = foreach A generate (chararray)$1 as res\_name, (chararray)$6 as team, (chararray)$9 as year, (chararray)$12 as sport, (chararray)$14 as medal;  Result = filter B by team == ‘India’ and medal == ‘Gold’;  dump Result;  output – we get total 138 rows. |
|  |  |