--explain.pig

divs = load 'NYSE\_dividends' as (exchange, symbol, date, dividends);

grpd = group divs by symbol;

avgdiv = foreach grpd generate group, AVG(divs.dividends);

store avgdiv into 'average\_dividend';

**Explain was written so that Pig developers could examine how Pig handled various scripts, thus its output is not the most user-friendly.**



**After optimizing the logical plan, Pig produces a physical plan. This plan describes the**

**physical operators Pig will use to execute the script, without reference to how they will**

**be executed in MapReduce.**



**Finally, Pig takes the physical plan and decides how it will place its operators into one**

**or more MapReduce jobs.**



**illustrate takes a sample of your data and runs it through your script, but as it encounters**

**operators that remove data (such as filter, join, etc.), it makes sure that some records**

**pass through the operator and some do not.**

divs = load 'NYSE\_dividends' as (e:chararray, s:chararray, d:chararray, div:float);

recent = filter divs by d > '2009-01-01';

trimmd = foreach recent generate s, div;

grpd = group trimmd by s;

avgdiv = foreach grpd generate group, AVG(trimmd.div);

illustrate avgdiv;

