

To-do

Problem Statement:

What are the movie titles that user has rated?

```
A = load '/movies.csv' using PigStorage(',') as (movieid,title,genres);
B = filter A by (movieid NOT Null) AND (title NOT Null);
movies = foreach B generate movieid,title;
C = load '/ratings.csv' using PigStorage(',') as
(userid,movieid,rating:double,timestamp:long);
D = filter C by (userid NOT Null) AND (movieid NOT Null) AND (rating
NOT Null);;
ratings = foreach D generate userid,movieid,rating;
E= join movies by movieid,ratings by movieid;
F = foreach E generate $1 as title, $2 as userid, $4 as rating;
G = group F by userid;
Result = foreach G generate group, F.title;
store Result into '/movies_rating' using PigStorage(',');
```

How many times a movie has been rated by the user?

```
A = load '/movies.csv' using PigStorage(',') as (movieid,title,genres);
B = filter A by (movieid NOT Null) AND (title NOT Null);
movies = foreach B generate movieid,title;
C = load '/ratings.csv' using PigStorage(',') as
(userid,movieid,rating:double,timestamp:long);
D = filter C by (userid NOT Null) AND (movieid NOT Null) AND (rating
NOT Null);;
ratings = foreach D generate userid,movieid,rating;
E= join movies by movieid,ratings by movieid;
F = foreach E generate $1 as title, $2 as userid, $4 as rating;
G = group F by title;
Result = foreach G generate group, COUNT(F.rating);
store Result into '/movies_rating_result' using PigStorage(',');
```

What is the Average rating given for a movie?

```
A = load '/movies.csv' using PigStorage(',') as (movieid,title,genres);
B = filter A by (movieid NOT Null) AND (title NOT Null);
movies = foreach B generate movieid,title;
C = load '/ratings.csv' using PigStorage(',') as
(userid,movieid,rating:double,timestamp:long);
D = filter C by (userid NOT Null) AND (movieid NOT Null) AND (rating
NOT Null);
ratings = foreach D generate userid,movieid,rating;
E = join movies by movieid,ratings by movieid;
F = foreach E generate $1 as title, $2 as userid, $4 as rating;
G = group F by title;
Result = foreach G generate group, AVG(F.rating);
store Result into '/movies_rating_result' using PigStorage(',');
```