

To-do

Problem Statement:

What are the movie titles that the user has rated?

Pig Query:

```
movies = LOAD '/movies.csv' USING PigStorage(',') as (movieId,title,genres);
```

```
ratings = LOAD '/ratings.csv' USING PigStorage(',') as  
(userId,movieId,rating,timestamp);
```

```
Joinmoviesratings = join movies by movieId, ratings by movieId;
```

```
Result = ORDER Joinmoviesratings by ratings::rating;
```

```
Final_Result = limit Result 10;
```

ScreenShot Result:

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

Pig Query:

```
movies = LOAD '/movies.csv' USING PigStorage(',') as (movieId,title,genres);
```

```
ratings = LOAD '/ratings.csv' USING PigStorage(',') as (userId,movieId,rating,timestamp);
```

```
Joinmoviesratings = join movies by movieId, ratings by movieId;
```

```
C= group by rating;
```

```
E = foreach c generate group, COUNT(C.movieId);
```

```
F= order E by $1 DESC;
```

```
Result = limit F 10;
```

```
dump Result;
```

In question 2 above, what is the average rating given for a movie?

Pig Query:

```
movies = LOAD '/movies.csv' USING PigStorage(',') as (movieId,title,genres);
ratings = LOAD '/ratings.csv' USING PigStorage(',') as (userId,movieId,rating,timestamp);
Joinmoviesratings = join movies by movieId, ratings by movieId;

C= group by rating;

E = foreach c generate group, AVG(C.rating);

F= order E by $1 DESC;

Result = limit F 10;

dump Result;
```