Assignment 2 – DS8003

Amir Ghaderi – 500794236

October 13th, 2016

**Question 1:**

**HiveQL:** select userid, count(\*) as cnt from userratings group by userid order by cnt desc limit 1;

**Result:** 405 737

**Question 2:**

**HiveQL:** select avg(rating) as average from userratings where movieid ="178";

**Result:** 4.344

**Question 3:**

**HiveQL:** select occupation, count(rating) as cnt

from userinfo join userratings on

(userinfo.userid=userratings.userid)

group by occupation

order by cnt desc

limit 3;

**Result:** student 21957

other 10663

educator 9442

**Question 4:**

**HiveQL:** select count(distinct a.userid) from

userratings a join userinfo b on

(a.userid=b.userid)

where a.rating =5 and b.gender ="M";

**Result:** 657