QUERIES FOR EDX ANALYSIS

**CONTENTS**

1. **Queries**

**QUERIES FOR EDX ANALYSIS**

1. **Students who have registered for a course in their first login but have never activated their account**

mysql> select (select count(\*) from auth\_user where is\_active=0)/(select count(id) from auth\_user)\*100 as pct\_inactive;

1. **Dropouts (who never came back to continue the course)**

**2.1.**  Number of drop outs in the corresponding weeks.

(for the first week;;out of the total number of students):

mysql> select ((select count(\*) from auth\_user au where atediff(au.last\_login,au.date\_joined)<7)/(select count(\*) from (select distinct id from auth\_user) x))\*100 as pct\_dropouts;

**2.2.** Distrubution of drop outs according to their age group.

(for the age group 1975-1980;expressed in terms of percentage) :

mysql> select ((select count(\*) from auth\_user au,auth\_userprofile aup where datediff(au.last\_login,au.date\_joined)<7 and au.id=aup.user\_id and aup.year\_of\_birth>=1975 and aup.year\_of\_birth<1980)/(select count(\*) from (select au.id from auth\_user au,auth\_userprofile aup where au.id=aup.user\_id and aup.year\_of\_birth>=1975 and aup.year\_of\_birth<1980) x))\*100 as pct\_dropout\_age;

\*similarly for the other age groups\*

**2.3.** Distrubution of drop outs according to location.

Location column is emty.So we could not retrieve the data.

**2.4.** Distrubution of drop outs according to their education.

(expressed as total number of drop outs according to the groups):

mysql> select count(\*), level\_of\_education from (select \* from auth\_userprofile aup where aup.user\_id in (select au.id from auth\_user au where datediff(au.last\_login,au.date\_joined)<7))y group by y.level\_of\_education;

**2.5.** Distrubution of drop outs according to gender.

For males (out of the total number of males;expressed in terms of percentage):

mysql> select ((select count(\*) from auth\_user au,auth\_userprofile aup where datediff(au.last\_login,au.date\_joined)<7 and au.id=aup.user\_id and aup.gender='m')/(select count(\*) from (select distinct au.id from auth\_user au,auth\_userprofile aup where aup.gender='m' and au.id=aup.user\_id)x))\*100 as pct\_gender\_male;

For females (out of the total number of females;expressed in terms of percentage):

mysql> select ((select count(\*) from auth\_user au,auth\_userprofile aup

where datediff(au.last\_login,au.date\_joined)<7 and

au.id=aup.user\_id and aup.gender='f')/(select count(\*) from (select distinct au.id from auth\_user au,auth\_userprofile aup

where aup.gender='f' and au.id=aup.user\_id) x))\*100 as pct\_gender\_female;

**2.6.** Distrubution of drop outs according to course category.

(for the first course) :

mysql> select ((select count(\*) from auth\_user au,auth\_userprofile aup,courseware\_studentmodule csm where datediff(au.last\_login,au.date\_joined)<7 and au.id=aup.user\_id and csm.student\_id=au.id and csm.course\_id='Summer\_Intern\_IIT\_Mumbai /

SI001/2014\_SI\_May') / (select count(\*) from (select au.id from auth\_user au,auth\_userprofile aup,courseware\_studentmodule csm where au.id=aup.user\_id and csm.student\_id=au.id and csm.course\_id='Summer\_Intern\_IIT\_Mumbai/SI001/2014\_SI\_May') x))\*100 as pct\_dropout\_course\_category;

**2.7.** Distrubutions of drop outs according to combination of different factor mentioned as above.

**3. Dropouts (who came back to continue the course at the end: to note down the procrastinators):**

3.1 Distrubution of drop outs according to their age group.

3.2 Distrubution of drop outs according to location.

3.3 Distrubution of drop outs according to their education.

3.4 Distrubution of drop outs according to gender.

3.5 Distrubution of drop outs according to course category.

3.6 Distrubutino of drop outs according to combination of different factor mentioned as above.

**4.** **Grades of regular students VS grades of procrastinators.**

**5. Grades of students depending on the various factors:**

5.1 Distrubution of grades according to their age group.

5.2 Distrubution of grades according to location.

5.3 Distrubution of grades according to their education.

5.4 Distrubution of grades according to gender.

5.5 Distrubution of grades according to course category.

5.6 Distrubutino of grades according to combination of different factor mentioned as above.