Data Modeling and Databases: Assignment 6

Sokolov Maxim
email: somal1996@rambler.ru

Innopolis University

September 27, 2015

1 Question 1

Thinking: Based on uniform distribution we can imagine the following: all movies are shot all years. And in one year all movies are shot simultaneously. So all movies are shot part by part and year by year. So we can enter value w. W is part of movie are shot in one year. And w*40000=x, w*90=y, where x is amount of movies which are shot in one year and y is amount of movies which have one title (for example: Star Wars, Star Wars II, etc). And at the end w*90*40000=60000 (is amount of all movies)

(a) This query can be divided on two part: movies with "year" =1924 and movies with "year"=2001. But all years in our table between 1925 and 2014. So year $>$ 1924 and we must find count of movies with "year"=2001. In another words, we must find x and $x = 60000/90 = 666,6667$ (rounding).
Answer: 666,6667.

(b) Assume, A is set of Movies, where year=1999 and B is set of Movies where title is 'Fight Club'. Then answer is $|A \cup B|$. But $|A \cup B| = |A| + |B| - |A \cap B|$,where $|A|$ is x, $|B|$ is y, $|A \cap B|$ is. In the end $|A \cup B| = x + y - z = 60000/90 + 60000/40000 - 60000/(90 * 40000) = 668,1833$
Answer: $668,1833$.

(c) We have 5 variants of rating. And amount of needed values is 4 (2,3,4,5). Based uniform distribution, we can compute answer. It is $1300000 * 4/5 = 1040000$.
Answer: 1040000.

(d) This query shows count for each unique movie, so answer to this exercise is amount of distinct movies.
Answer: 40000.

(e) This query shows amount of distinct titles from Reviews and from Movies simultaneously. Amount of distinct imbd from Reviews is 33000 and amount of distinct imbd from Movies is 60000, so we have min(33000,60000)=33000 distinct imbd from both tables. But x is amount titles on one imbd, so amount of distinct titles is $33000/x = 22000$
Answer: 22000.

2 Question 2

1. (a)
   - yes, because we have following tuples (A,B): (α,9), (α,16).
   - no, because each value of B corresponds one value of B.
   - yes, because we have following tuples (C,A): (F,α), (F,β).
   - no, because each value of AC corresponds one value of B no, because each value of B corresponds one value of B.
   - no, because each value of B corresponds one value of AC.

   (b) no, we cannot identify FD based this tuples, because we not know relations between attributes and we can add tuples which can violate FD.

2. (a) b

   (b)
   - yes, P→Q,Q→R ⇒ P→R.
   - yes, PS→P,P→Q,PS→TRV⇒PS→QT;PS→QT→UR→U.
   - no, because for U we must have P, but we can not create P
   - no, because we can not get P

   (c) False, because $\{Q\}^+$=$\{Q,R\}$.
   {Q}:Q→R
   {Q,R}: end

   (d) True, $\{P,S\}^+$=$\{P, Q, R, S, T, U, V\}$.
   {P,S}: P→Q.
   {P,S,Q}: Q→R.
   {P,S,Q,R}: PS→TRV.
   {P,S,Q,R,T,V}:QT→UR.
   {P,S,Q,R,T,V,U}:end.