```
Q1
 SELECT TOP 5 TagName,Count FROM Tags ORDER BY Tags.Count DESC ;
 SELECT COUNT(*) FROM post;
 SELECT COUNT(*) From tagonly WHERE Tag1='r' or Tag2='r' or Tag1='regression'
 or Tag2='regression' or Tag3='regression' or Tag4='regression'
 or Tag1='time-series' or Tag2='time-series' or Tag3='time-series'
 or Tag1='machine-learning' or Tag2='machine-learning' or Tag3='machine-learning'
 or Tag1='probability' or Tag2='probability' or Tag3='probability' or Tag4='probability';

 Q2
 SELECT AVG(Score) FROM post WHERE post.PostTypeId=1;
 SELECT AVG(Score) FROM post WHERE post.PostTypeId=2;


Q3
CREATE TABLE pearson (TotalScore int, PostCount int,Id int,Reputation int);
INSERT INTO pearson
SELECT SUM(post.Score),COUNT(users.Id),users.id,users.Reputation
FROM post, users WHERE users.Id=post.OUI GROUP BY users.Id,users.Reputation;
## in R
pearson<-read.csv("file:///Users/zhangyuting/Desktop/xml/pearson.csv")
cor(pearson[,c("TotalScore","Reputation")],method="pearson")


Q4
SELECT AVG(users.UV)
FROM post, users WHERE users.Id=post.OUI and post.PostTypeId=1;
SELECT AVG(users.UV)
FROM post, users WHERE users.Id=post.OUI and post.PostTypeId=2;

Q5
 CREATE TABLE answer(answerId int, ADate datetime,AOUI int);
INSERT INTO answer
 SELECT post.Id,post.CD,OUI FROM post
 WHERE  PostTypeId=2;
  CREATE TABLE question(questionId int, AAId int,QDate datetime,QOUI int);
INSERT INTO question
 SELECT post.Id,post.AAI,post.CD,OUI FROM post
 WHERE  PostTypeId=1 and AAI is not NULL;
 SELECT questionId,AAId,answerId,ADate,QDate,
 DATEDIFF(HOUR,QDate,ADate) AS response
 FROM answer JOIN question On AAId=answerId ORDER BY response DESC;


Q6
 SELECT users.Id, DATEDIFF(HOUR,users.CD,ADate) AS ATime,
  DATEDIFF(HOUR,users.CD,QDate) AS QTime,
   DATEDIFF(HOUR,users.CD,comments.Date) AS CTime
 FROM answer,users,question,comments
 WHERE AOUI=users.Id and QOUI=users.Id and comments.Id=users.Id;
### save query as  “sequense.csv” and write codes in R

### Q6 second part using R
```