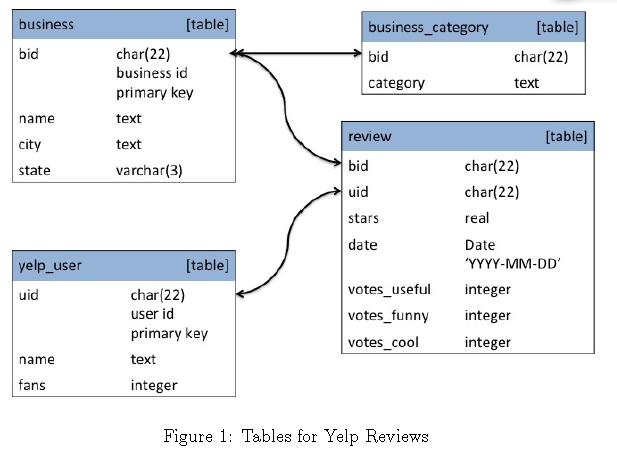
SQL Midterm Exam：Yelp Reviews

For this question you will look into Yelp reviews collected from multiple cities across the world. Figure 1 gives the schema of the tables you will use.



For example, the tuple in the review table

(“1202”, “1001”, 3.0, “2011-08-10”, 2, 1, 0)

means that the #1202 business was reviewed by #1001 user on August 10th, 2010 and got 3 stars. 2 other users think this review is useful, 1 user thinks this review is funny and 0 users think this review is cool.

Questions:

1. Count the distinct categories of businesses.
2. Count the number of businesses in Pennsylvania (“PA”).
3. Find every business in Pennsylvania that has the word “Coffee” (case-sensitive) in its name but is not classified as a coffee place (i.e. has no word “Coffee" in any of its categories). List the business' bid (i.e. business id) and name in ascending bid order.
4. List Pennsylvania businesses that have more than 10 reviews and all of them are “5 stars”. (This is probably a sign of fraud). Order by the number of reviews received (decreasing), and bid (i.e. business id, increasing). For each business, print bid, name and number of reviews.
5. We refer to users that have only voted once as one-hit wonders and such users are likely to be fake accounts. In this query, find the businesses that got more than 150 5-star reviews from such one-hit wonders. For each business, print its bid (i.e. business id), name and the number of reviews from one-hit wonders. Order by the number of reviews received from one-hit wonders (decreasing), then business name (increasing) then bid (increasing).