|  |  |  |
| --- | --- | --- |
| **1a** | The goal of this problem is to generate a lot of non-repeating data from a relatively small collection. With this in mind, you are to develop a script whereby you can generate a minimum of 5 million (5,000,000) unique rows of test data for a table containing data on the contact person for each of our customers. Name the table: **HW1\_Q1** **\_CONTACT**  The table should have the following attributes (columns):   * First Name * Last Name * City * State   You may take your initial names and locations from Eagle Database or any other source that you find, but you must document where you take it from and how you use it.  The data used to populate the table should be appropriate to the column, NOT gibberish or random numbers. For example: city should be ‘Lafayette’, ‘Greenville’, ‘Riverdale’ etc., NOT ‘asdf;lkjadsf’ or ‘23adsf898’.  Your answer should include all DDL and DML statements in the specific sequence necessary to generate and populate the **HW2\_Q1** **\_CONTACT** table. Include sufficient directions to the grader (in comment statements) to allow him/her to run your test-data generation process.  **Notes:**   * if your **HW1\_Q1** **\_CONTACT** table contains less than 100,000 rows of unique data you will receive NO POINTS (e.g., no partial credit) for this question! * If your **HW1\_Q1\_CONTACT** table was generated from hard coding, you will receive a 20% deduction of the maximum available points. * Use of a 3rd party tool to generate your data is NOT acceptable. You are fully equipped at this point in the semester to generate this data using SQL and/or PL/SQL. | 12 pts. |
| **1b** | Write one or more queries to prove that the table you created in 1a has at least 5 million rows of unique data. Show the number of unique First Names, unique Last Names, unique States and unique Cities.  Include both the queries and your results in the submission. | 3 pts. |