Stored Procedures Assignment 2

Create the database `sp2` by downloading the file `sp2.sql` and using the `File>Run SQL Script` menu command. This database has two tables: `wm_dcs` and `wm_stores`. The first table gives information about Wal-Mart distribution centers for dry goods in the United States. The second database contains data for all Wal-Mart stores and Sam's Club locations in the United States. The first table has these data fields:

- `dc_id`: four-digit identifier for each DC
- `address`: street address
- `city`: city
- `state`: two-letter state abbreviation
- `zip`: zip code, sometimes with four-digit extension, which is represented as `VARCHAR()` data
- `lat`: latitude (in degrees)
- `lon`: longitude (in degrees)

The second table has these data fields:

- `store_id`: unique identifier for each store
- `name`: name of store location from web source
- `lat`: latitude (in degrees)
- `lon`: longitude (in degrees)

Wal-Mart frequently realigns which distribution centers serve each of its stores as the number of stores increase, as new distribution centers are built, as distribution strategies change, and as the sales volume of individual stores grow at differing rates.

A preliminary step in formulating an optimization model for determining the supply chain structure would be to reduce to the extent possible the number of DCs that would be considered as a possible supply for each store: a majority of DCs should not be considered to supply any particular store because they are so much farther away than the closest alternatives. By eliminating these DC-Store combinations from consideration in the optimization model the number of decision variables can be significantly reduced so that the model runs faster. One criterion that can be used for culling out the feasible DCs for each store is to consider Wal-Mart's policy of having each store be within one day's drive for its truck drivers. This policy makes it easier for Wal-Mart to recruit drivers because it can advertise that drivers can "be home every night." Recruiting a sufficient number of truck drivers is always a difficult task so this policy makes sense both for the consideration of the drivers and for the success of the business. This policy effectively limits the distance between any DC and the stores that it supplies to 250 miles. A small number of stores, however, may not have a DC within 250 miles in which case they should be served by the closest DC or either of the closest two DCs, considering DC capacity limitations and other constraints.

Your task is to write a stored procedure named `spGetCloseDCsStores` that returns these fields, `dc_id`, `store_id`, `distance`, for the DCs that fit these criteria for each Store:

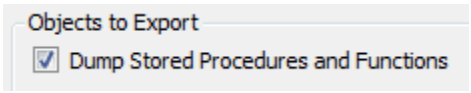
- The closest four DCs within 250 miles of each store or as many DCs (less than four) that are within 250 miles of each store, ***or***
- If there are no DCs within 250 miles of a store, then return the closest two DCs.

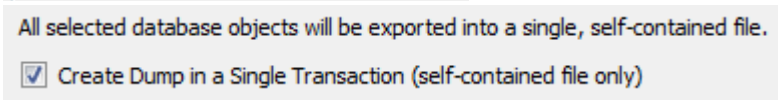
Every store should have at least one DC associated with it in the returned data with these criteria.

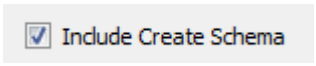
An empty stored procedure named `spGetCloseDCsStores` is included in the `sp2` database to facilitate naming standardization across all assignments and the checking of homework: please write your stored procedure using it and the `Alter Stored Procedure` feature which is available by right-clicking on the stored procedure to alter it and add your code. Besides returning `dc_id`, `store_id`, and `distance` in a result set from your stored procedure, also use `SELECT` statements to return statistics about your results including:

- The total number of DC-Store combinations you generated using the criteria above
- For each DC provide a `COUNT()` of the number of stores associated with it in the results you obtain

When your assignment is complete, create a "dump" file of your database that includes your stored procedure for the task described above and submit it to Blackboard. Choose all three of these options when creating your dump file:

- 

- 

- 

The stored procedure is a preliminary step in the optimization of Wal-Mart's supply chain structure. *We will not be optimizing Wal-Mart's supply chain in this assignment: we are only preparing some of the data that would go into defining the decisions variables and the objective function for an optimization model.*

There are several ways that distance between DCs and Stores could be computed. One is to rely on an API, such as Google Maps. In this exercise we will be approximating the distance "as the crow flies" and ignoring the degree to which the roads from a DC to a store are direct, or not. Thus, we will be relying on an equation, specifically the Haversine formula, to estimate distance, which is a function of latitude and longitude of the locations (in units of radians). Here are two links that describe the formula:

- http://andrew.hedges.name/experiments/haversine/
- http://www.movable-type.co.uk/scripts/latlong.html

Latitude and longitude are most often expressed in degrees and, when this is the case, those data must be converted to radians to implement the Haversine formula. That conversion is accomplished by multiplying latitude and longitude in degrees by π /180, which is due to a circle having 360 degrees which is equivalent to 2π radians. The database `sp2`, which you will upload into Workbench using the script `sp2.sql`, has a stored procedure template, `spHaversineExample`, with an example applying the Haversine formula using these variables:

- `storeLat`: Latitude of store location in degrees
- `storeLon`: Longitude of store location in degrees
- `dcLat`: Latitude of DC location in degrees
- `dcLon`: Longitude of DC location in degrees
- `storeLatRad`: Latitude of store location in radians
- `storeLonRad`: Longitude of store location in radians
- `dcLatRad`: Latitude of DC location in radians
- `dcLonRad`: Longitude of DC location in radians

Using this equation requires that you successfully create latitudes and longitudes in radians from the latitude and longitude data in the database (in degrees) and embed this equation in a computed field in a query, or however you might want to use it.

Notes on Data Acquisition

Reading this section is not required for this assignment. However, in case you are interested, it described where the data for this assignment came from.

- The Wal-Mart store locations came from this site:
  - https://fusiontables.google.com/DataSource?docid=1ag3Z3Uwp_hWiHeiBRqGrS_HzEtwUjeVh4d4ZAnI#rows:id=1
- Wal-Mart published its list of regional distribution centers here:
  - http://www.mwpvl.com/html/walmart.html
- The previous source listed street address, which were converted to latitude and longitude using this site:
  - http://www.latlong.net/convert-address-to-lat-long.html