Assignment 10: Databases and Data Structures

Date Due: December 9, 2013 Instructor: Trani

Problem

Download the structured file called "amtrakStations.mat" from our site. This struct file contains all US Amtrak train stations. Each station has the following information:

```
% Sample structured file contains Amtrak station data
% File name: amtrakStations
%   Geometry   - GIS type of data (points or polygons)
%   X          - Longitude (negative is East of London)
%   Y          - Latitude (positive is North of the Equator)
%   STNCODE    - Station code
%   STNNAME    - Station name
%   ADDRESS1
%   ADDRESS2
%   CITY       - station city
%   STATE      - state where station is located
%   ZIP        - zip code
%   STNTYPE    - station type
%   STFIP      - station FIP number (A unique number)
```

Task 1:

Create a Matlab script that loads the amtrakStation data file. The script should also load the US map provided and used in class. Plot all Amtrak train stations and superimpose the US map. Make the x and y labels clearly visible with font size 20. Your stations should be shown with red circles in the map. Your map should have a grid.

Task 2:

Improve the Matlab script created in Task 1. This part of the analysis implement a FOR loop to identify select the train stations located in the state of Virginia and plots these stations with a black triangle as marker. Other stations are shown in red circles as in Task 1. Display in the command window the number of stations found in Virginia.

Hint: you can use the Matlab string comparison function (strcmp) to compare the state name ('VA') with the "children" of amtrakStations called amtrakStations.STATE. The "child" amtrakStations.STATE contains the name of the US state where the station is located.

Task 3:

Improve the Matlab script created in Tasks 1 and 2. This part of the analysis implement another FOR loop to identify and select the stations located the city of San Francisco and plot these stations with a blue triangle as marker. Other stations are shown in red circles as in Task 1. Display in the command window the number of stations found in San Francisco.

Problem 2

A data file called damData2.xlsx contains data for 8,130 dams in the US. The database has been collected from the National Transportation Atlas Database (http://www.rita.dot.gov/bts/sites/rita.dot.gov.bts/files/publications/national_transportation_atlas_database/2013/points.html).

Dam No.	Name	DamID	Longitude	Latitude	County	River	City	Owner	Owner Type
1	BRONCE	PR00027	-66.5616	18.0173	PONCE	CHANNEL FROM LAGO GILES	COTO LAUREL	SUCESION J. S	P
2	COAMO	PR00016	-66.3844	18.0179	COAMO	COAMO	FELICIA	COMMONWEA	S
3	PATILLAS	PR00023	-66.021	18.021	PATILLAS	RIO GRANDE DE PATILLAS	PATILLAS	COMMONWEA	S
4	LOCO	PR00004	-66.8863	18.0446	YAUCO	LOCO	PALOMAS	COMMONWEA	S
5	ANA MARIA 5	PR00031	-66.5616	18.0583	PONCE	INABON RIVER OFFSTREAM	COTO LAUREL	SUCESION J. S	P
6	PORTUGUES	PR82202	-66.635	18.07	NONE	PORTUGUES	PONCE	CESAJ	F
7	CARITE	PR00021	-66.1066	18.0782	GUAYAMA	LA PLATA	GUAYAMA	PREPA (FORM	U
8	ANTONIO LUCCHETTI	PR00003	-66.865	18.0831	YAUCO	YAUCO	YAUCO	PREPA (FORM	U
9	GUAYABAL	PR00013	-66.5033	18.0888	JUANA DIAZ	JACAGUAS	JUANA DIAZ	COMMONWEA	S
10	TOA VACA DAM	PR00014	-66.485	18.1033	VILLALBA	TOA VACA	JUANA DIAZ	PR AQUEDUC	U

Dam Type	Purpose	Year	Height	Max Storae	Normal Storage	Surface Area	Drain Area	Hazard	State
RE	I	1939	52	843	620	0	0	S	PR
CBRE	I	1914	65	1280	0	145	66	H	PR
RE	I	1914	147	13797	11029	367	25	H	PR
CNPG	IS	1951	74	2059	1039	69	8	H	PR
RE	I	1939	53	2382	1927	79	0	H	PR
CNVA	CSR		274	32000	16841	230	10	H	PR
RE	HSIR	1913	104	14992	8953	333	8	H	PR
CNPG	HIS	1952	175	17595	11575	266	17	H	PR
CNCB	IS	1913	130	5933	4768	373	21	H	PR
REER	IS	1972	215	54875	50620	836	22	H	PR

Nomenclature

The official National Inventory of Dams identification number for the dam, known formerly as the National Id. The first two characters of the identity are the State two letter abbreviation, based on the location of the dam. The last five characters of the identity are a unique number (AB#####).

DAM_TYPE

CB = Buttress

CN = Concrete

ER = Rockfill

MS = Masonry

MV = Multi-arch

OT = Other

PG = Gravity

RE = Earth

ST = Stone

TC = Timber Crib

VA = Arch

PURPOSES

A code or codes indicating the purposes for which reservoir is used, in order of importance. The codes are concatenated if the dam has multiple purposes. For example, SCR would indicate the primary purposes of Water Supply and Flood Control and Storm Water Management, followed by Recreation. The data may contain words or abbreviations that were used instead of the appropriate codes; these have been retained.

YR_COMPLETED

The year when the main dam structure was completed. A value of 0 or 9999 indicates that the year is unknown.

HEIGHT

The maximum of either the dam height (the vertical distance between the lowest point on the crest of the dam and the lowest point in the original streambed), the hydraulic height (the vertical distance between the maximum designed water level and the lowest point in the original streambed), or the structural height (the vertical distance between the lowest point of the excavated foundation and the top of the dam) of the dam. The height is given in feet, to the nearest foot.

MAX_STOR

The maximum storage in acre-feet. Maximum storage is the total storage space in a reservoir below the maximum attainable water surface elevation, including any surcharge storage.

NORMAL_STO

The normal storage in acre-feet. Normal storage is the total storage space in a reservoir below the normal retention level, including dead and inactive storage and excluding any flood control surcharge storage. A value of 0 may indicate that the normal storage is unknown or may indicate that the dam is normally dry.

SURF_AREA

The surface area, in acres, of the impoundment at its normal retention level.

DRAIN_AREA

The drainage area of the dam in square miles. It is defined as the area that drains to a particular point (in this case, the dam) on a river or stream.

HAZARD

A term indicating the potential hazard to the downstream area resulting from failure or mis-operation of the dam or facilities. The potential hazard is low (L). A dam where failure or misoperation results in no probable loss of human life and low economic and/or environmental loss. Losses are principally limited to the owner's property.

The potential hazard is significant (S). A dam where failure or misoperation results in no probable loss of human life but can cause economic loss, environmental damage, disruption of lifeline facilities, or impact other concerns. These dams are often located in predominantly rural or agricultural areas but could be located in areas with population and significant infrastructure.

SYMBOL

A code indicating how the dam is symbolized in the National Atlas Map Maker. The symbol is determined from the primary Purpose code given for the dam, which is the first or only code listed under the attribute Purpose. Where the Purpose attribute contains a word or abbreviation instead of the appropriate code, the first letter of the Purpose value was still used to determine the symbol code. In these cases the symbol may not accurately represent the true primary purpose of the dam.

The dam or reservoir contained by the dam has a primary purpose other than flood control and storm water management; debris control; fish and wildlife pond; hydroelectric; irrigation; navigation; fire protection, stock, or small farm pond; recreation; water supply; or tailings.

Task 1:

Import the data using the Matlab import wizard. Save the variables as individual vectors with names presented in row 1 of the Excel file. Save the information into a Matlab binary file for later tasks.

Task 2:

Create a Matlab script to retrieve the binary file created in Task 1. This file contains all the information about dams in the US. Include code that will plot the locations of all dams in the US. For this task you will use reuse the US map file employed in Problem 1.

Task 3:

Improve the Matlab script of Tasks 2 to search for all dams made of concrete. Plot the concrete dams in a US map.

Task 4:

Improve the Matlab script of Tasks 3 to search for all dams made built before 1895. Plot two sets of dams in a US map: a) those constructed before 1930 and b) those constructed after 1930.Use different colors and symbols to differentiate the dams. Zoom

into the area of Virginia and comment on the number of dams built before 1930. Display (in the command window) the number of dams built before and after 1930.

Problem 3

Download the two Excel files named: a) virginia_airports.xls and b) runways_virginia_airports.xls provided in the syllabus web site. Samples of the files are shown below for reference.

Table 2. Sample File virginia_airports.xls.

Name	ID	Ownership	Airport_ID	Latitude	Longitude
VIRGINIA HIGHLANDS	VJI	PU	25600.*A	36.6871111	-82.033333
ALPHA NATURAL RESOURCES	22VG	PR	25600.01*H	36.6966667	-81.995333
MOUNT VERNON HOSPITAL	VA82	PR	25609.*H	38.7403922	-77.0772
ALEXANDRIA HOSPITAL HELISTOP	9VA2	PR	25609.01*H	38.8226128	-77.104145
HILL TOP	VA64	PR	25617.*A	37.4159831	-77.95389
MERLIN AERODROME	2VA3	PR	25617.2*A	37.315	-77.866111
TIMBERDOODLE	93VA	PR	25618.*A	37.5362533	-79.023356

Table 3. Sample File runways_virginia_airports.xls.

Airport_ID	Runway Label	Length (feet)	Width (ft)	Surface	Lights
25600.*A	06/24	4471	75	ASPH-G	
25600.*A	06/24	4471	75	ASPH-G	ODALS
25600.01*H	H1	600	100	TURF	
25600.01*H	H1	600	100	TURF	
25609.*H	H1	75	75	CONC	
25609.*H	H1	75	75	CONC	

Task 1:

Import both files into Access creating a new database. In the virginia_airports.xls file **define the Airport_ID filed as the primary** key. Define Airport_ID as a text datatype in both tables. Establish a one-to-many relationship between the virginia_airport and the runways_virginia_airports files using the field Airport_ID. Verify that the relationship works. Explain how do you know the relationship works.

Task 2:

Create an MS Access query to find all the airports in Virginia whose runways longer than 5,300 feet. State the number of airports and show the resulting Access table. In your solution table include the following fields:

Airport name

Airport ID

Runway label

Runway length

Task 3:

Create another MS Access query to find the **public airports** in Virginia (Ownership field = PU) with **runways with lengths between 3,000 and 7,500 feet** and with **Asphalt (ASPH) runways**. State the number of airports and show the resulting Access table showing the following fields:

Airport name

Airport ownership

Airport ID

Runway label

Runway length

Task 4:

Create an MS Access query to find all the runways at airports in Virginia with **Medium-intensity Approach Lighting System with Sequenced Flashing Lights (MALSF**) or **Medium-intensity Approach Lighting System with Runway Alignment Lights (MALSR**). The query should produce a table with the airport name, runway label, runway length, lights system name, airport latitude and longitude. A picture of a MALSF system deployed at Atlanta International Airport is shown below.

Figure 1. MALSF Runway Approach Light System (A.A. Trani).

Task 5:

Export the results of the query performed in Task 4 to a text or Excel file. Read the exported file into Matlab and make a plot of the airport locations with MALSF light systems using the provided map of Virginia provided.