REVIEW QUESTIONS

1. Give an example of data exploration from your own experience.
2. Download % population change by county for your state between 1990 and 2000. (You can get the data from the Census Bureau's website, **http://www.census.gov/,** or from the GIS data clearinghouse website in your state.) Use the data to compute the median, first quartile, third quartile, mean, and standard deviation.
3. Use the county data and descriptive statistics from Question 2 to draw a boxplot. What kind of data distribution does the boxplot show?
4. Among the graphics presented in Section 10.1.2, which are designed for multivariate (i.e., two or more variables) visualization?
5. Figure 10.4 exhibits a weak positive relationship between % population change, 1990–2000 and % persons under 18 years old, 2000. What does a positive relationship mean in this case?
6. Describe brushing as a technique for data exploration.
7. Describe an example of using spatial aggregation for data exploration.
8. Describe an example of using map comparison for data exploration.
9. Refer to Figure 10.14, and write an SQL statement to query the owner name of parcel P104.
10. Refer to Figure 10.14, and write an SQL statement to query the owner name of parcel P103 OR parcel P104.
11. Refer to Table 10.1, and fill in the blank for each of the following query operations:

    [Create a new selection] "cost" > 8

    _____ of 10 records selected

    [Add to current selection] "soiltype" = 'N3' OR "soiltype" = 'Ns1'

    _____ of 10 records selected

    [Select from current selection] "area" > 400

    _____ of 10 records selected

    [Switch Selection]

    _____ of 10 records selected
12. Refer to Box 10.5, and describe an example of using "intersect" for spatial data query.
13. Refer to Box 10.5, and describe an example of using "are contained by" for spatial data query.
14. You are given two digital maps of New York City: one shows landmarks, and the other shows restaurants. One of the attributes of the restaurant layer lists the type of food (e.g., Japanese, Italian, etc.). Suppose you want to find a Japanese restaurant within 2 miles of Times Square. Describe the steps you will follow to complete the task.
15. Can you think of another solution for Question 14?
16. Refer to Figure 10.19. If the query statement is ([slope] = 1) AND ([aspect] = 3), how many cells in the output will have the cell value of 1?

APPLICATIONS: DATA EXPLORATION

This applications section covers six tasks. Task 1 uses the select feature by location tool. Task 2 lets you use the chart engine, introduced in ArcGIS 9.2, to create a scatterplot and link the plot to a table and a map. In Task 3, you will query a joint attribute table, examine the query results using a magnifier window, and bookmark the area with the query results. Task 4 covers relational database query. Task 5 combines spatial and attribute data queries. Task 6 deals with raster data query.

Task 1: Select Feature by Location

What you need: *idcities.shp,* with 654 places in Idaho; and *snowsite.shp,* with 206 snow courses in Idaho and the surrounding states.

Task 1 lets you use the select feature by location method to select snow courses within 40 miles of Sun Valley, Idaho, and plot the snow station data in charts.

1. Start ArcCatalog, and connect to the Chapter 10 database. Launch ArcMap. Add *idcities.shp* and *snowsite.shp* to Layers. Right-click Layers and select Properties. On the General tab, rename the data frame Tasks 1&2 and select Miles from the Display dropdown list.
2. Step 2 selects Sun Valley from *idcities.* Choose Select By Attributes from the Selection menu. Select *idcities* from the layer dropdown list and "Create a new selection" from the method list. Then enter in the expression box the following SQL statement: "CITY_NAME" = 'Sun Valley'. (You can click Get Unique Values to get Sun Valley from the list.) Click Apply and close the dialog. Sun Valley is now highlighted in the map.
3. Choose Select By Location from the Selection menu. In the Select By Location dialog, choose "select features from" from the first dropdown list, check *snowsite,* choose "are within a distance of" from the second list, choose *idcities* from the third list, enter 40 miles for the distance to buffer, and click Apply. Snow courses that are within 40 miles of Sun Valley are highlighted in the map.
4. Right-click *snowsite* and select Open Attribute Table. Click Selected to show only the selected snow courses.

Q1. How many snow courses are within 40 miles of Sun Valley?

5. ArcGIS 9.2 introduced a couple of new features for managing tables and selected records. First, the Options menu has the print table option. Second, double marking allows the user to select and highlight records (and features) from the currently selected records. For example, Vienna Mine Pillow has the highest SWE_MAX among the selected records. To see where Vienna Mine Pillow is located on the map, you can click on the far left column of its record. Both the record and the point feature are highlighted in yellow.
6. Leave the table and the selected records open for Task 2.

Task 2: Make Dynamic Chart

What you need: *idcities.shp* and *snowsite.shp,* as in Task 1.

Task 2 lets you create a scatterplot from the selected records in Task 1 and take advantage of a live connection among the plot, the attribute table, and the map.

1. Make sure that the *snowsite* attribute table shows the selected records from Task 1. This step exports the selected snow courses to a new shapefile. Right-click *snowsite* in Tasks 1&2, point to Data, and select Export Data. Save the output shapefile as *svstations* in the Chapter 10 workspace. Add *svstations* to Tasks 1&2. Select Clear Selected Features from the Selection menu for *snowsite,* and close the *snowsite* attribute table.
2. Next, create a chart from *svstations.* Click the Tools menu in ArcMap, point to Graphs, and select Create. In the first panel of the Create Graph Wizard, select ScatterPlot for the graph type, *svstations* for the layer/table, ELEV for the Y field, and SWE_MAX for the X field. Click Next. In the second panel, enter Elev-SweMax for the title. Click Finish. A scatterplot of ELEV against SWE_MAX appears.

Q2. Describe the relationship between ELEV and SWE_MAX.

3. The scatterplot is dynamically linked to the *svstations* attribute table and the map. Open the *svstations* attribute table, if necessary. Click a point in the scatterplot. The point, as

well as its corresponding record and feature, is highlighted. You can also use the mouse pointer to select two or more points within a rectangle in the scatterplot. This kind of interaction can also be initiated from either the attribute table or the map.

4. Right-click the scatterplot. The contact menu offers various options for the plot, such as print, save, export, and add to layout.

Task 3: Query Attribute Data from a Joint Attribute Table

What you need: *wp.shp,* a timber-stand shapefile; and *wpdata.dbf,* a dBASE file containing stand data.

Data query can be approached from either attribute data or spatial data. Task 3 focuses on attribute data query.

1. Insert a new data frame in ArcMap and rename it Task 3. Add *wp.shp* and *wpdata.dbf* to Task 3. Next join *wpdata* to *wp* by using ID as the common field. Right-click *wp,* point to Joins and Relates, and select Join. In the Join Data dialog, opt to join attributes from a table, select ID for the field in the layer, select *wpdata* for the table, select ID for the field in the table, and click OK.

2. *wpdata* is now joined to the *wp* attribute table. Open the attribute table of *wp.* The table now has two sets of attributes. Click the Options dropdown arrow and choose Select By Attributes. In the Select By Attributes dialog, make sure that the method is to create a new selection. Then enter the following SQL statement in the expression box: "wpdata.ORIGIN" > 0 AND "wpdata.ORIGIN" <= 1900. Click Apply.

Q3. How many records are selected?

3. Click Selected at the bottom of the table so that only the selected records are shown. Polygons of the selected records are also highlighted in the *wp* layer. To narrow the selected records, again choose Select By Attributes from the Options dropdown menu. In the Select By Attributes dialog, make sure that the method is to select from current selection. Then prepare an SQL statement in the expression box that reads: "wpdata.ELEV" <= 30. Click Apply.

Q4. How many records are in the subset?

4. To take a closer look at the selected polygons in the map, click the Window menu in ArcMap and select Magnifier. When the magnifier window appears, click the window's title bar, drag the window over the map, and release the title bar to see a magnified view.

5. Before moving to the next part of the task, select Clear Selection from the Options menu in the *wp* attribute table and click All to show all records. Then choose Select By Attributes from the same menu. Enter the following SQL statement in the expression box: ("wpdata.ORIGIN" > 0 AND "wpdata.ORIGIN" <= 1900) AND "wpdata.ELEV" > 40. (The pair of parentheses is for clarity; it is not necessary to have them.) Four records are selected. The selected polygons are all near the top of the map. Zoom to the selected polygons. You can bookmark the zoom-in area for future reference. Click the Bookmarks menu, and select Create. Enter *protect* for the Bookmark Name. To view the zoom-in area next time, click the Bookmarks menu, and select *protect.*

6. ArcGIS 9.2 added the MyPlaces function, which can save bookmarks to a file. For example, to save the four polygons from Step 5 to MyPlaces, you can select MyPlaces from the Tools menu. In the next dialog, choose Select Feature(s) from the Add From dropdown menu. The dialog now shows four places, one for each polygon. You can click on one of the places and click Zoom To or Pan To to look at the place. The Save button on the MyPlaces dialog allows you to save the four places to a place file for future use.

Task 4: Query Attribute Data from a Relational Database

What you need: *mosoils.shp,* a soil map shapefile; *component.dbf, coeplants.dbf,* and *comonth.dbf,* three dBASE files derived from the SSURGO database developed by the Natural Resources Conservation Service (NRCS).

Task 4 lets you work with the SSURGO database. By linking the tables in the database properly, you can explore many soil attributes in the database from any table. And, because the tables are linked to the soil map, you can also see where selected records are located.

1. Insert a new data frame in ArcMap and rename it Task 4. Add *mosoils.shp, component.dbf, coeplants.dbf,* and *comonth.dbf* to Task 4.
2. First, relate *mosoils* to *component.* Right-click *mosoils* in the table of contents, point to Joins and Relates, and click Relate. In the Relate dialog, select mukey from the first dropdown list, select *component* from the second list, select mukey from the third list, enter soil_comp for the relate name, and click OK.
3. Next prepare two other relates: comp_plant, relating *component* to *coeplants* by using cokey as the common field; and comp_month, relating *component* to *comonth* by using cokey as the common field.
4. The four tables (the *mosoils* attribute table, *component, coeplants,* and *comonth*) are now related in pairs by three relates. Right-click *comonth* and select Open. Click the Options dropdown arrow and choose Select By Attributes. In the next dialog, create a new selection by entering the following SQL statement in the expression box: "flodfreqcl" = 'Frequent' OR "flodfreqcl" = 'Occasional'. Click Apply. Click Selected at the bottom of the table so that only the selected records are shown.

Q5. How many records are selected in *comonth?*

5. To see which records in *component* are related to the selected records in *comonth,* go through the following steps: Click the Options dropdown arrow in the *comonth* table, point to Related Tables, and click comp_month: component. The *component* table appears with the related records. You can find which records in *coeplants* are related to those records that have frequent or occasional annual flooding by using comp_plant: coeplants with the *component* table.
6. To see which polygons in *mosoils* are subject to frequent or occasional flooding, do the following: Click the Options dropdown arrow in the *component* table, point to Related Tables, and click soil_comp: mosoils. The attribute table of *mosoils* appears with the related records. And the *mosoils* map shows where those selected records are located.

Q6. How many polygons in *mosoils.shp* have a plant species with the common plant name of "Idaho fescue"?

Task 5: Combine Spatial and Attribute Data Queries

What you need: *thermal.shp,* a shapefile with 899 thermal wells and springs; *idroads.shp,* showing major roads in Idaho.

Task 5 assumes that you are asked by a company to locate potential sites for a hot-spring resort in Idaho. You are given two criteria for selecting potential sites:

- The site must be within 2 miles of a major road.
- The temperature of the water must be greater than 60°C.

The field TYPE in *thermal.shp* uses *s* to denote springs and *w* to denote wells. The field TEMP shows the water temperature in °C.

1. Insert a new data frame in ArcMap. Add *thermal.shp* and *idroads.shp* to the new data frame. Right-click the new data frame and select Properties. On the General tab, rename

the data frame Task 5 and choose Miles from the Display dropdown list.

2. First select thermal springs and wells that are within 2 miles of major roads. Choose Select By Location from the Selection menu. Do the following in the Select By Location dialog: choose "select features from" from the first dropdown list, check *thermal,* select "are within a distance of" from the second list, select *idroads* from the third list, and enter 2 (miles) for the distance to buffer. Click Apply. Thermal springs and wells that are within 2 miles of roads are highlighted in the map.

Q7. How many thermal springs and wells are selected?

3. Next, narrow the selection of map features by using the second criterion. Choose Select By Attributes from the Selection menu. Select *thermal* from the Layer dropdown list and "Select from current selection" from the Method list. Then enter the following SQL statement in the expression box: "TYPE" = 's' AND "TEMP" > 60. Click Apply.

4. Open the attribute table of *thermal.* Click Selected at the bottom of the attribute table so that only the selected records are shown. The selected records all have TYPE of *s* and TEMP above 60.

5. Map tips are useful for examining the water temperature of the selected hot springs. Right-click *thermal* in the table of contents and select Properties. On the Display tab, check the box to Show Map Tips (uses primary display field). On the Fields tab, select TEMP from the Primary display field dropdown list. Click OK to dismiss the Properties dialog. Click Select Elements on the standard toolbar. Move the mouse pointer to a highlighted hot-spring location, and a map tip will display the water temperature of the spring.

Q8. How many hot wells and springs are within 5 kilometers of *idroads* and have temperatures above 70?

Task 6: Query Raster Data

What you need: *slope_gd,* a slope raster; and *aspect_gd,* an aspect raster.

Task 6 shows you different methods for querying a single raster or multiple rasters.

1. Select Data Frame from the Insert menu in ArcMap. Rename the new data frame Task 6, and add *slope_gd* and *aspect_gd* to Task 6.

2. Select Extension from the Tools menu and make sure that Spatial Analyst is checked. Click the View menu, point to Toolbars, and check Spatial Analyst. The Spatial Analyst toolbar appears. Select Raster Calculator from the Spatial Analyst menu. In the Raster Calculator dialog, prepare the following statement in the expression box: [slope_gd] = 2. (The = sign appears as = = in the expression box.) Click Evaluate. The layer *Calculation* is added to the table of contents. Cells with the value of 1 are areas with slopes between 10 and 20 degrees.

Q9. How many cells in *Calculation* have the cell value of 1?

3. Select Raster Calculator from the Spatial Analyst menu, and prepare the following statement in the expression box: [slope_gd] = 2 AND [aspect_gd] = 4. (The word AND is shown as & in the expression box.) Click Evaluate. Cells with the value of 1 in *Calculation2* are areas with slopes between 10 and 20 degrees and the south aspect.

Q10. What percentage of area covered by the above two rasters has [slope_gd] = 3 AND [aspect_gd] = 3?

Challenge Task

What you need: *cities.shp,* a shapefile with 194 cities in Idaho; *counties.shp,* a county shapefile of Idaho; and *idroads.shp,* same as Task 5.

cities.shp has an attribute called CityChange, which shows the rate of population change between 1990 and 2000. *counties.shp* has attributes

on 1990 county population (pop1990) and 2000 county population (pop2000). Add a new field to *counties.shp* and name the new field CoChange. Calculate the field values of CoChange by using the following expression: ([pop2000] − [pop1990]) × 100/[pop1990]. CoChange therefore shows the rate of population change between 1990 and 2000 at the county level.

Q1. What is the average rate of population change for cities that are located within 50 miles of Boise?

Q2. How many counties that intersect an interstate highway have CoChange >= 30?

Q3. How many cities with CityChange >= 50 are located within counties with CoChange >= 30?

REFERENCES

Andrienko, G., N. Andrienko, P. Jankowski, D. Keim, M.-J. Kraak, A. MacEachren, and S. Wrobel. 2007. Geovisual Analytics for Spatial Decision Support: Setting the Research Agenda. *International Journal of Geographical Information Science* 21: 839–57.

Andrienko, N., G. Andrienko, A. Savinov, H. Voss, and D. Wettschereck. 2001. Exploratory Analysis of Spatial Data Using Interactive Maps and Data Mining. *Cartography and Geographic Information Science* 28: 151–65.

Andrienko, N., G. Andrienko, H. Voss, F. Bernardo, J. Hipolito, and U. Kretchmer. 2002. Testing the Usability of Interactive Maps in CommonGIS. *Cartography and Geographic Information Science* 29: 325–42.

Anselin, L. 1999. Interactive Techniques and Exploratory Spatial Data Analysis. In P. A. Longley, M. F. Goodchild, D. J. Maguire, and D. W. Rhind, eds., *Geographical Information Systems,* 2d ed., pp. 253–66. New York: Wiley.

Batty, M., and Y. Xie. 1994. Modelling Inside GIS: Part I. Model Structures, Exploratory Spatial Data Analysis and Aggregation. *International Journal of Geographical Information Systems* 8: 291–307.

Becker, R. A., and W. S. Cleveland. 1987. Brushing Scatterplots. *Technometrics* 29: 127–42.

Buja, A., D. Cook, and D. F. Swayne. 1996. Interactive High-Dimensional Data Visualization. *Journal of Computational and Graphical Statistics* 5: 78–99.

Cleveland, W. S. 1993. *Visualizing Data.* Summit, NJ: Hobart Press.

Cleveland, W. S., and M. E. McGill, eds. 1988. *Dynamic Graphics for Statistics.* Belmont, CA: Wadsworth.

Crampton, J. W. 2002. Interactivity Types in Geographic Visualization. *Cartography and Geographic Information Science* 29: 85–98.

DiBiase, D., A. M. MacEachren, J. B. Krygier, and C. Reeves. 1992. Animation and the Role of Map Design in Scientific Visualization. *Cartography and Geographic Information Systems* 19: 201–14, 265–66.

Dykes, J., A. M. MacEachren, and M.–J. Kraak (eds.). 2005. *Exploring Geovisualization.* Amsterdam: Elsevier.

Egbert, S. L., and T. A. Slocum. 1992. EXPLOREMAP: An Exploration System for Choropleth Maps. *Annals of th Association of American Geographers* 82: 275–88.

Haining, R. 2003. *Spatial Data Analysis: Theory and Practice* Cambridge, England: Cambric University Press.

MacEachren, A. M., F. P. Boscoe, D. Haug, and L. W. Pickle. 19 Geographic Visualization: Designing Manipulable Maps Exploring Temporally Varyin Georeferenced Statistics. *Proceedings of the IEEE Symposium on Information Visualization,* pp. 87–94.

Meyer, M. A., F. R. Broome, and R. H. J. Schweitzer. 1975. Co Statistical Mapping by the U. Bureau of the Census. *Americ Cartographer* 2: 100–17.

Monmonier, M. 1989. Geographi Brushing: Enhancing Exploratory Analysis of the Scatterplot Matrix. *Geograph Analysis* 21: 81–84.

Olson, J. 1981. Spectrally Encod Two-Variable Maps. *Annals o the Association of American Geographers* 71: 259–76.

Peterson, M. P. 1995. *Interactive and Animated Cartography.* Englewood Cliffs, NJ: Prentice Hall.

Scott, L. M. 1994. Identification GIS Attribute Error Using