

MSIA 430 : Introduction to Software Tools for BI**Spring 2022****Assignment-1 (problems for Part II only)****Due: Monday, May 19 at 11:59PM**

Purpose: This assignment is intended to provide an opportunity to demonstrate your understanding of the two major portions of the course so far: GraphDB and Spatial Data management. It consists of two parts:

1. The problem-set posted for you as a “Rolling-Assignment-1”, for which no submission was required at the time. At this point, you need to provide the solutions to those problems.
2. The two problems below, pertaining to the Spatial Data management part of the course.

NOTE: this is a team assignment, and one submission per team is sufficient – however, you need to make sure to include the names all the team members in the submission. Typed solutions are strongly preferred (no illegible handwriting, please – if you have to, make sure that whatever portion is handwritten in your solutions, it is done in a clear and legible manner...).

What to turn-in: Combine the solutions for your “rolling assignment” and the solutions for the two problems below, and upload them in the corresponding rubrique on Canvas.

Problem 1: Consider a database equipped with a Spatial Extender, for which you are given the following DB schema:

Person(pID, p-address, p-city, p-zip, age)
Company(cID, c-address, c-city, c-zip, founding-year)
Works-for(pID, cID, salary)
Stores(sID, storeType, s-address, s-city, s-zip)
Building(bID, useType, b-address, b-city, b-zip, floorPlan)
Citites(cityName, cityBoundary, population)
RoadNet(rID, category, segments)

You can assume that the corresponding types for the spatial attributes are “as-usual” (e.g., floorPlan: Polygon; cityBoundary: Polygon; segments: LineString) – i.e., the types from SQL/MM Spatial. You can also assume that all the corresponding operators and functions (both topological and geometrical) discussed in class to be available.

Write the SQL statement for the following queries:

Q1: Retrieve all the employees (i.e., persons working for a company) who are older than 35 and live in the same city as the company that they work for, and for whom the distance from their residential building to the work-place building is less than 5 miles.

Q2: Retrieve all the roads of category ‘highway’ that are intersecting cities for which the total area occupied by the stores (in each respective city) is greater than 5 mi².

Problem 2: Consider the scenario shown in Fig. 1 (below), illustrating a set of ten 2-D points with their coordinates.

- (a) Illustrate the structure of the *space-driven* Quadtree that will index the set of points in Fig. 1, where the boundary of the region of interest is a square, with lower-left and upper-right end points at (0,0) and (200, 200). Assume that the capacity of every node is 1 – i.e., it can only store coordinate values for 1 point (plus, of course, the pointers/references to its children).
- (b) Consider the query: *Retrieve all the points inside the (axes-parallel) rectangle R with lower-left corner at (12,8) and upper-right corner at (55,120)*. For the tree structure from part (a), show all the nodes in the Quadtree that will be visited (as well as pruned) in an attempt to answer the query.

Figure 1. Set of points' distribution

