



# Module 5

## Extended Query Formulation with SQL

### Lesson 1: Query Formulation Guidelines

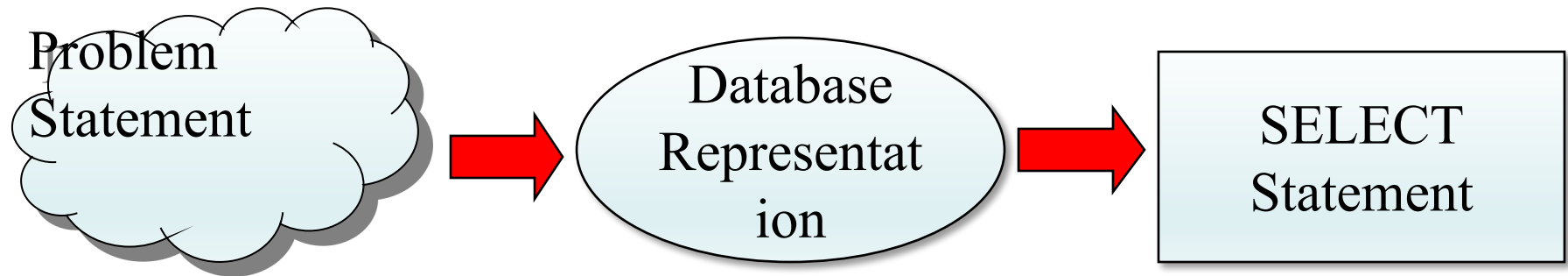


# Lesson Objectives

- Convert a problem statement into a database representation using the critical questions
- Identify extra tables in a SELECT statement



# Query Formulation Process

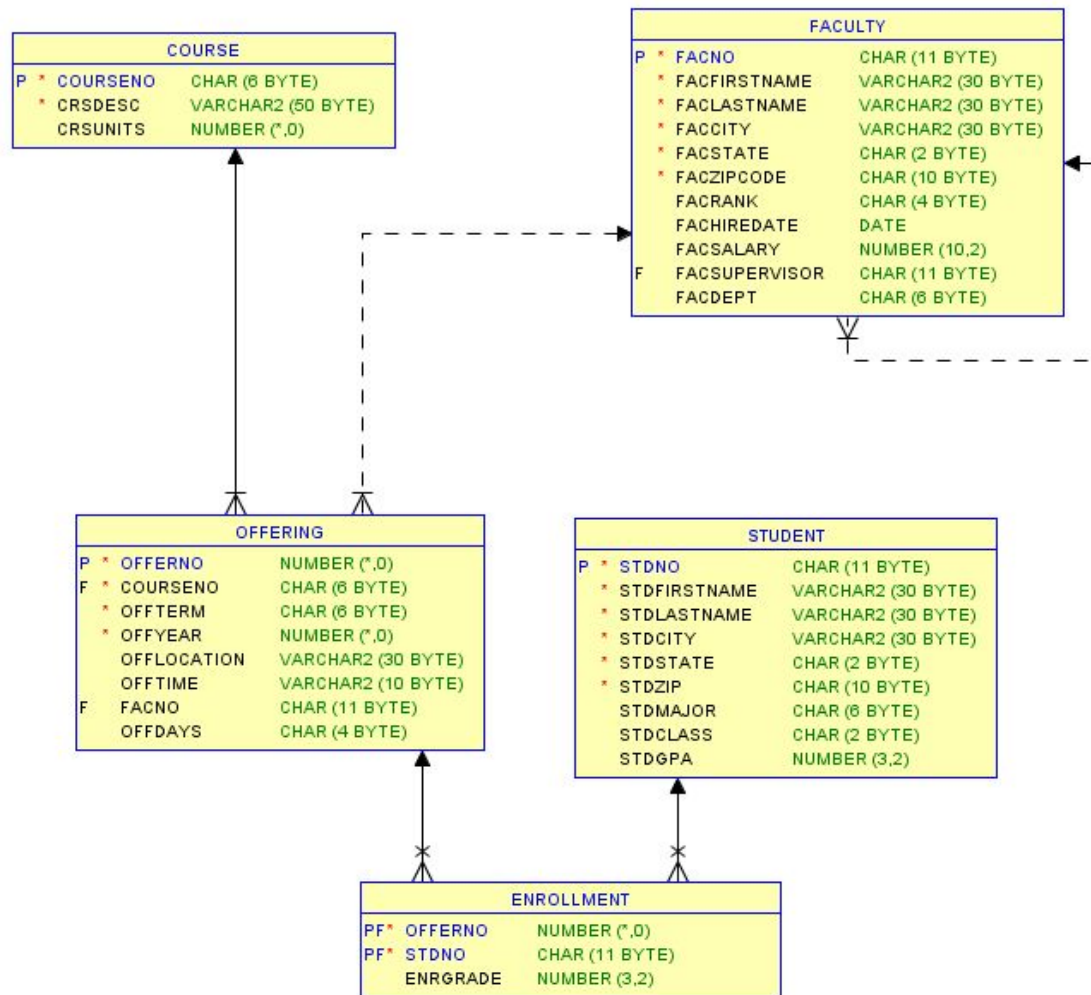


# Critical Questions

- What tables?
  - Columns in result
  - Conditions to test (including join conditions)
- How to combine the tables?
  - Usually join of PK to FK
  - More complex ways to combine
- Individual rows or groups of rows?
  - Aggregate functions in result
  - Conditions with aggregate functions



# University Database Diagram



# Summarization and Joins I

Example 1: List the number of students enrolled in each 2017 course offering showing the offer number and number of students in the result.

```
SELECT Offering.OfferNo,  
       COUNT(*) AS NumStudents  
FROM Enrollment, Offering  
WHERE Offering.OfferNo = Enrollment.OfferNo  
      AND OffYear = 2017  
GROUP BY Offering.OfferNo;
```



# Summarization and Joins II

Example 2: List the offering number, course number, and average GPA. Only include courses offered in fall 2016 in which the average GPA is greater than 3.0.

```
SELECT Enrollment.OfferNo, CourseNo,  
        AVG(StdGPA) AS AvgGPA  
FROM Offering, Enrollment, Student  
WHERE Offering.OfferNo = Enrollment.OfferNo  
      AND Student.StdNo = Enrollment.StdNo  
      AND OffYear = 2016 AND OffTerm = 'FALL'  
GROUP BY Enrollment.OfferNo, CourseNo  
HAVING AVG(StdGPA) > 3.0;
```





# Efficiency Considerations

- Little concern for efficiency
- Intelligent SQL compilers
- Correct and non redundant solution
  - No extra tables
  - No unnecessary grouping
  - No missing join conditions





# Extra Table Redundancy

Example 3: List the offering number, course number, and average GPA. Only include courses offered in fall 2016 in which the average GPA is greater than 3.0.

```
SELECT Enrollment.OfferNo, Offering.CourseNo,  
        AVG(StdGPA) AS AvgGPA  
FROM Offering, Enrollment, Student, Course  
WHERE Offering.OfferNo = Enrollment.OfferNo  
      AND Student.StdNo = Enrollment.StdNo  
      AND Course.CourseNo = Offering.CourseNo  
      AND OffYear = 2016 AND OffTerm = 'FALL'  
GROUP BY Enrollment.OfferNo, Offering.CourseNo  
HAVING AVG(StdGPA) > 3.0;
```



# Summary

- Remember the query formulation process
- Use critical questions to convert a problem statement into a database representation
- Check for unnecessary tables and missing join conditions
- Much practice with more difficult problems involving joins and grouping



# Database Representation to SQL Statement

