

## **Introduction to Data Management and Engineering**

### **1<sup>st</sup> Assignment**

**Q1.** I used inner join to combine the employee table with department table with their common variable. Then I kept the rows where the department number includes employees with salary greater than 42000. Lastly, grouped by department name, I counted the number of each employee in each department that fulfilled the necessary criterion.

**Q2.** I found the dno in which the employee with the minimum salary belongs. Then, for these specific dnos, I selected the names of the employees. The minimum salary was 25000 and people from different departments had that salary, so I picked all the departments that fulfilled that criterion.

**Q3.** First, I picked the dnos where the dname was research. Then for these dnos I calculated the average salary. Lastly, I showed the names which had a salary bigger than the average plus 5000.

**Q4.** I joined the table employee with table department based on their common variable and then I kept those rows where mgrssn equals ssn, which corresponds to an employee being a manager.

**Q5.** First, I computed the number of employees in each department (counted the names grouped by the dno). Then, I computed the number of projects in each department (I counted the project numbers grouped by dnum). Next, I inner joined them together and with department table to obtain the department names and lastly, I inner joined them with the view of question 4 to obtain the manager names.

**Q6.** First, I inner joined works\_on table with employee table to obtain the number of males and females for each project (grouped by pno). Then, I inner joined the previous with a table that brings the sum of workhours in each project and the number of employees in each project. Next, I inner joined all that with department table to obtain the dname and then with project table to obtain the projects.

**Q7.** First, I created a table with the sum of the hours for each employee (grouped by essn) and I ranked them with descending order. Rank assigns the rank number to each row in a partition and

skips the number for similar values. The result brought multiple values for rank 2, as they had the same value. Also, I dropped the null values for working hours. Then, I inner joined all that with employee table to get the names, next with department table to get department names and lastly with works\_on table to get the hours for each specific project. Lastly, I kept the rows which had a rank  $\leq 3$ . As I mentioned before, the values were more than 3 and the second rank had multiple same values.

**Q8.** I created a table which obtain the table teaches and inner joined it with instructor table to get the instructor names. Then, I inner joined it with takes table using course\_id, sec\_id, semester, year to get the grades. Next, I kept only the grades that were not null and grouped all of them by the name of instructor while calculating the number of A,B,C,F (count when grade equals A etc) that each instructor gave. Lastly, I ordered everything by the number of As, then by Bs, Cs and Fs.

**Q9.** First, I created a table that brings the total hours in a semester for each time slot (grouped by time\_slot\_id) and inner joined it with the table section based on the time\_slot\_id. Then I created a table, that brings the total number of students grouped by course, sec\_id, semester and year and I inner joined it with the all the previous based on course, sec\_id, semester and year.

**Q10.** In this question I inner joined section table with time\_slot table to obtain the days and starting hours, then with course table to obtain title and department number, then with teaches table to obtain sec\_id, semester, year using all the common variables and lastly with instructor table to obtain the name of instructor. Finally, I chose to order them first by the year and then by semester, sec\_id, day (M=1,... R=4,F=5) and starting hour. Also, I chose to show the ending hour of the course, the sec id, the course\_id and department name because I believe that the calendar will be displayed in a more coherent manner.