Warm up (5 points)
1. (5 points) Write your name and student-id on your blue book.

Logical data design (10 points)
2. (10 points) The following ER diagram shows a person must have at most one spouse. Reduce this ER diagram to minimal number of tables.

![ER Diagram]

Structured Query Language, SQL (30 points)
3. Recall tables that describe sailors who reserve boats at different times:
   - Sailors (sid, sname, rating)
   - Boats (bid, color, bname)
   - Reserve (sid, bid, date)

   a. (10 points) Write the SQL command to delete those sailors with a rating below average.
   b. (10 points) Write the SQL command to retrieve the name of those sailors with the maximum rating.
   c. (10 points) For each red boat, retrieve its id and the average rating of sailors who have reserved this boat.

Relational algebra (30 points)
4. (15 points) Using the tables specified for Question number 3, write the algebraic expression to retrieve id of those sailors who have reserved both a red and a green colored boat.
5. (15 points) Write the algebraic expression to retrieve the color of boats with at least one reservation.

Indexes and simple query processing (25 points)
Assume 10 sailor records fit per disk page. Assuming the sailor table has 100,000 records, it consists of 10,000 disk pages.
6. (10 points) With a clustered hash index on the rating attribute of sailor table (see Question 3 for the Sailor schema), what is an estimated number of disk I/Os required to process the following exact-match selection predicate: “sname=John”. Assume the hash index consists of 20,000 buckets and the records are uniformly distributed across the buckets. Explain your answer.
7. (15 points) Assuming a 2-level deep non-clustered B+-tree on the rating attribute of sailor table, what is an estimated number of disk I/Os required to retrieve those
sailors with a rating higher than seven: “rating>7”. You may assume the minimum rating for a sailor is 1 and the maximum is 10. You may also assume the rating is an integer value with records uniformly distributed across the possible values. Explain your answer.