

## CS430/630 – Homework 2

50 points (5/100 of final grade)

**Instructions:** The homework is **due at 19:00 EST on Thu March 18**. The submission must have **TWO TEXT FILES attached** called **Q1.sql** and **Q2.sql**. The submission email **MUST BE SENT FROM YOUR @UMB ACCOUNT** and must have subject formatted **EXACTLY AS BELOW**:

**[CS430] HW2**                    - for CS430 students, and

**[CS630] HW2**                    - for CS630 students.

The timestamp will be considered according to the UMB email server. Submissions received (i.e., tagged) later than 19:00 EST on March 18 **WILL NOT BE GRADED**. You are only allowed a **SINGLE** submission (i.e., you cannot send multiple emails – only the **first received** will be graded).

**Email submissions must be sent to Nicholas J Pankewytch at <N.Pankewytch001@umb.edu>**

All questions have equal weight.

**Tips:** 1.) Use sqlplus to test out your queries on sample data you come up with. This will help you to understand how sqlplus works and how to modify your queries to make them work properly.

2.) USE COMMENTS. You are more likely to get partial credit if you make it clear what you are trying to accomplish rather than just writing the query with no context.

### Question 1 (25 points)

You are given the following relational schema.

Books(bid:integer, bname:string, author:string, year:integer, price:integer)

Orders(cid:integer, bid:integer, quantity:integer)

Customers(cid:integer, cname:string, zipcode:string)

The meaning of attributes is as follows:

bid: book unique identifier, bname: book name; author, year and price (per unit) are self-explanatory

quantity: number of books purchased with an order

cid: unique customer identifier, cname: customer name, zipcode: customer address zipcode.

Write **SQL queries** for the following:

- Find distinct authors of books ordered in quantities of at least 50 from zipcode '02125'.
- Find the names of customers who ordered only books with price lower than \$100.
- Find the average price of ordered books for each customer (listed as cid) who bought at least 20 different book titles costing at least \$20 each.
- Find the names of customers who ordered every book written by 'Edgar Codd'.
- Find the author name(s) of the best-selling book(s) (meaning the books that sold the highest number of copies).
- [630 only] Find the name(s) of most valuable customer(s), i.e. those who spent the most money.

## Question 2 (25 points)

You are given the following relational schema.

```
Passengers(pid:integer, pname:string, age:integer, city:string)
Flights(fid:integer, from:string, to:string, miles:integer, aircraft:string)
Tickets(pid:integer, fid:integer, price:integer)
```

The database contains passenger and flight information for an airline in a single day. The description is as follows: passengers are uniquely identified by `pid` and they have as attributes name, age and city of residence. Flights are uniquely identified by `fid`. Each flight has an origin and destination airport identified by `from` and `to` respectively (airport codes are 3-letter strings, and are unique – no two airports have the same code - you do not need to enforce that constraint, just assume it is respected), flight distance given by `miles`, and `aircraft` type, e.g., B737, B777, B787, A320, etc. The `Tickets` relation provides information of flights that passengers took, and also the `price` that was paid for that flight **segment** (assume that each ticket corresponds to a segment; for people with connecting flights, they will have separate tickets, billed independently).

Write **SQL queries** for the following:

- Find the ages of passengers who traveled only on B787 aircraft (assumes they were on at least one flight).
- Find the name of passengers who traveled on all aircraft types.
- For each flight with at least 100 passengers, find the average price paid by passengers who are older than 30. List the `fid` in the output to identify the flight.
- Calculate the revenue earned by the airline from tickets on the flight(s) with the longest distance.
- For every origin airport, list the total number of passengers that departed from that airport, the total dollar amount earned by the airport for tickets on flights that depart from that airport, and the average distance for flights departing from that airport.
- [630 only] Find the age(s) of the passenger(s) who collected the most reward miles (assume that a reward mile is earned for each traveled mile).