

# CS348: Introduction to Database Systems

(Fall 2018)

## Assignment 4 (due at noon on Tuesday, November 20th)

**Overview:** This assignment consists of three questions. As an aid to scheduling your work on this assignment, you should plan on spending three to 6 hours total on the questions.

**Assignment submission:** A PDF file “a4.pdf” containing your answers to all questions:

**Page 1:** an ER diagram, the first part of Question 1,

**Page 2:** the second part of Question 1,

**Pages 3, 4 and 5:** SQL DDL defining the relational schema, the first part of Question 2,

**Pages 6 and 7:** the second part of Question 2, and

**Page 8:** Question 3.

### Question 1.

Assume your company is developing a digital camera online purchasing system for sale to camera stores. An initial analysis phase of the project has resulted in the following informal description of relevant data for the system.

- A store will be selling a variety of digital cameras and lenses. The digital cameras can have a combination of the following features.
  1. Cameras with a built-in lens.
  2. Cameras with an ability to replace lenses.
  3. Cameras with an electronic viewfinder.
  4. Cameras with an optical viewfinder.
  5. Cameras with a “through the lens” optical viewfinder.
  6. Cameras with an optical rangefinder.

Note that any combination of these features is possible with two exceptions: no camera will have more than one feature from the set {4, 5, 6}, and every camera must have exactly one feature from the set {1, 2}.

- Properties of all cameras that are relevant include the manufacturer, model number, date of product release, sensor size, pixel number, retail price and the number currently in stock.
- A property of cameras with a built-in lens that is relevant is an aperture range.
- Cameras with an ability to replace lenses are related to at least two or more lenses.
- Properties of a lens (not built-in to a camera) that are relevant include the manufacturer, model number, date of product release, aperture range, retail price and the number currently in stock.
- A lens, either built-in to a camera or not, is either a prime lens or a telescopic lens. In the former case, it has a relevant focal length, and in the latter case, a relevant focal length range.
- Online customers are either domestic customers or foreign customers.
- Properties of customers that are relevant include a unique customer number, a customer name, an email address and a shipping address.
- Each customer has any number of purchase orders, including possibly none at all. A subset of the purchase orders are in the process of being prepared for shipment and are therefore outstanding.
- Each purchase order is for either a camera or a lens, and will also have a selling price.
- Each camera or lens will have at least one customer evaluation.
- A customer evaluation is given by an individual customer and consists of a score between 1 and 5 (from bad to good) and a customer comment.

Do each of the following two parts:

1. Specify a conceptual design with an ER diagram that is capable of storing such information that formalizes as much of the informal description as possible.
2. Clarify any parts of the above informal description that are not captured by your ER diagram.

**Question 2.**

Do each of the following two parts:

1. Translate the ER diagram you produced for the previous question into SQL DDL commands that define a relational schema. The commands should include primary and foreign key constraints where appropriate.
2. Write two relational algebra queries on your relational schema that identify cameras or lenses that violate either of the following two integrity constraints:
  - At least one of the lenses that is related to a camera with an ability to replace lenses is a prime lens.
  - The retail price of any prime lens is higher than the retail price of any other prime lens with the same focal length whenever the lowest value of the aperture range of the latter is higher than the lowest value of the aperture range of the former.

**Question 3.**

Assume a relation  $R$  has four attributes  $\{A, B, C, D\}$ . For each of the following sets of FDs, assuming in each case that the dependencies in that set hold for  $R$ , do the following: (a) identify the candidate key(s) for  $R$ , and (b) identify the best normal form among  $\{1NF, 3NF, BCNF\}$  that  $R$  satisfies. If the best is 1NF then show why 3NF is violated, and if the best is 3NF then show why BCNF is violated.

1.  $\{ABC \rightarrow D, D \rightarrow A\}$
2.  $\{A \rightarrow B, BC \rightarrow D, A \rightarrow C\}$
3.  $\{AB \rightarrow C, AB \rightarrow D, C \rightarrow A, D \rightarrow B\}$