CPSC 368: Databases in Data Science
Creating an ER Diagram

Assignment Description

Based on the feedback received from your brainstorming assignment, pick a topic for your project. This is the topic you will focus on from now until the end of the semester so be sure to choose carefully.

After choosing a topic, design an ER diagram to represent the information from your data files.

We expect each ER diagram to have:

- At least 5 different entities and 5 relationships, not counting weak entities and ISAs
    - All entities must have their keys clearly identified using the notation discussed in lecture
    - An ISA relationship only counts as one entity. E.g., a parent entity with five children only counts as one entity, not six.
    - The parent of a weak entity counts as an entity but the weak entity itself does not.
- At least one meaningful ISA or non-trivial weak entity.
- For each relationship, identify the cardinality constraint and other constraints, such as participation constraints. We encourage you to provide a mix of constraints (some 1-to-many, many-to-many, 1-to-1).

Deliverables

All of the following items must be put together into a **single** PDF file.

1. A statement clearly identifying your project topic and what datasets you are using. There should be links to the datasets you wish to use.

2. An ER diagram that follows the requirements listed above. It is OK to hand-draw it but if it is illegible or messy or confusing, marks will be taken off.
    a. You can use software to draw your diagram (e.g., draw.io, GoogleDraw, Microsoft Visio, Powerpoint, Gliffy, etc.) The result should be a **legible** PDF or PNG document. Note that your ER diagram must use the conventions from the textbook and the lectures. For example, **do** not use crow's feet notation or notation from other textbooks).
    b. Please limit your diagram to a letter size page (8.5 x 11 inches). If you require additional space, talk to us **beforehand**.

3. A discussion about the various design choices you had to make as you created the ER diagram. Specifically discuss situations where you had to choose between representing

something as 1) an entity vs. an attribute, 2) an entity vs. a relationship, and 3) a binary vs. a ternary relationship OR an aggregated vs. non-aggregated relationship.

When discussing each design choice, you should be specific about which part of the ER diagram you are referring to and explain what led you to choose one design over another. You should also discuss pros/cons between the two designs.