Lab 1
Due in Gradescope (Lab Assignment #1)

Names: Brennan R. Nick R. Joey K.

Instructions

- Answer each of the questions below as a group in one document
- Only one submission per group
- *At the end of your document provide a division of labor*
    - Must be decided unanimously
    - Use percentages totaling 100%
- Upload to gradescope for “Lab Assignment #1” as a PDF
- Add all group members to submission

Answer the following questions.

1. What is metadata?

- A description of a database including data structures, types, and constraints.

2. Explain the difference between a database schema (or intension) and a database state (or extension).

- The database schema (intension) does not change very frequently, it consists of how the database is built (structure, data types, constraints)
- The database state is the actual data stored in the database at a specific time, this includes a collection of all the data, it changes every time the database is updated. Also known as the database instance
    a. Give an example of a database schema (list the relations and attributes and key attributes; we will talk about foreign keys later so don’t worry about them yet)
    - Example: University

- Entities: STUDENT, COURSES, SECTION, DEPARTMENT, INSTRUCTORS
- Student has a class
- Student has a section
- Student has a teacher
- Student has a key ID_number
- Student has a dorm
- Instructor has students
- Instructor has a course
- Courses have prerequisite courses
- Courses have a key ID_number
- Instructor has a key ID_number

b. Give an example of a database state using the schema in a. (A small number of records in each table is okay)

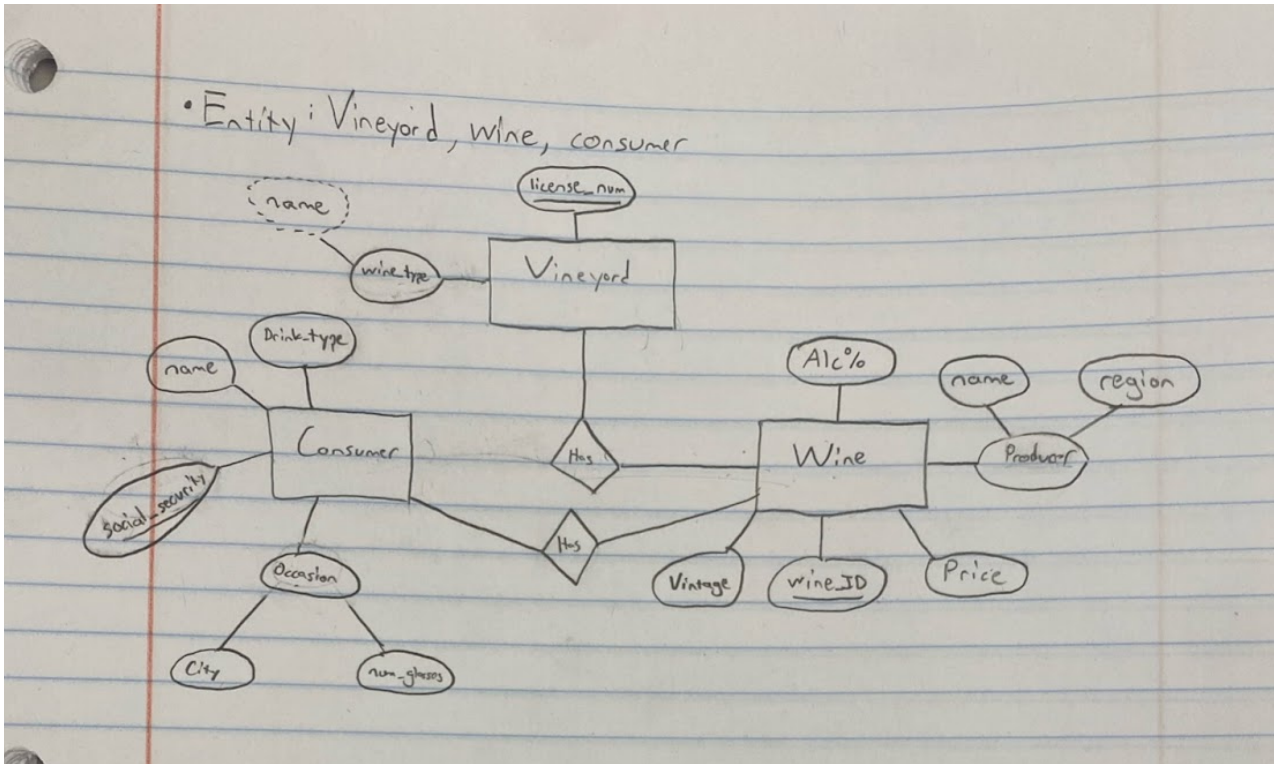
- When a new semester starts, the database is in an initial state because the student picks new courses, has no grades, and does not have an assigned meal plan. During the semester, the database is at a valid state because students have passed the drop date for classes and they are fully enrolled.

3. This example is from Valduriez and Gardarin – Analysis and Comparison of Relational Database Systems, Addison-Wesley, 1989.

Imagine that a winery needs to keep track of their vineyard productions and customers. They collect data about the wines, the producers of the wines, and even the drinking behaviors of their customers for better planning of their business. So let's look at the data that they are collecting. Here is a list of the items of data:

- Vineyards (by name such as Julienas, Medoc, etc.) for particular wines (W1, W2, …)
- Vintage of the wine (year)
- Percentage alcohol (usually in the range of 8-15%)
- Price of a bottle of the wine (in range of $0 to $100)
- Producer of a wine, including name and region]

- Name of a drinker of wine, and the type of drinker (light, average, or heavy)]
- Quantity of a particular wine harvested and the producer who harvested it
- List of known occasions where wine is drunk by a particular drinker including location (city), quantity (number of glasses)
    a. Given this description of data, develop a conceptual model of this data set that you could map to a relational database. Use the Entity-Relationship Model.



Be sure to include the following:

- Primary key attributes (Sometimes you need to add a data item to serve as a unique ID if there is not already one supplied in the original data description)
- Relationships
- Cardinality constraints
- Participation constraints

Do not develop tables yet; this is the conceptual model. We will learn to map to tables later.

b. For each of your named relationships, specify their type (specialization/generalization, association, or aggregation).

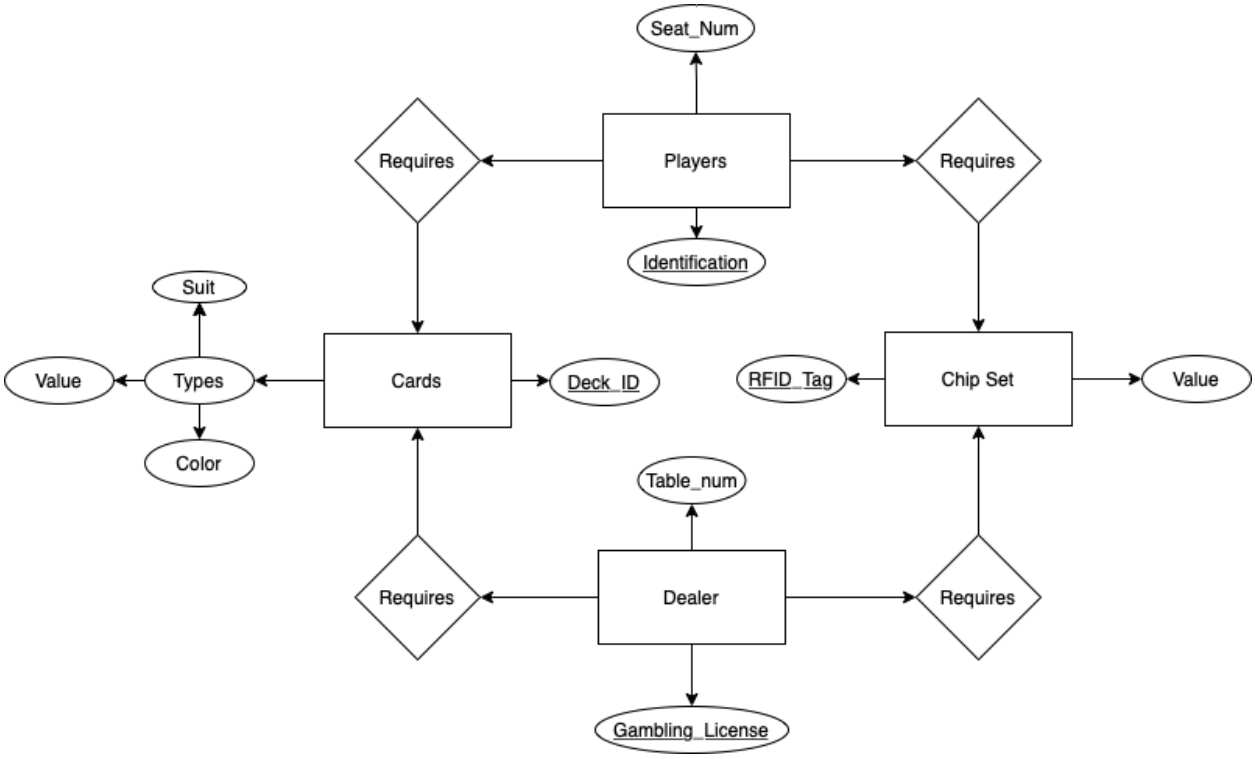
- The relationship between the vineyard and the wine is specialization. This is because the vineyard can produce many types of wine and have many different types of consumers.
- The consumer and wine have an aggregation relationship because if either/or are removed from the database, they will continue to exist independently.

c. Imagine that there are sociologists/psychologists, wine growers, and marketers who each wish to use this database. Can you construct one view of the data that you think each might be interested in? Describe the view in words, including the attributes from the data that would be included.

- These users would be interested in the name of a drinker of wine, and the type of drinker as well as the List of known occasions where wine is drunk by a particular drinker including location (city), quantity (number of glasses).
- This information could be compiled to find how likely the person is to drink based on the data and where they are most likely to drink.

d. Develop another E-R diagram of a data set in which you are interested. For example, you might tap into your interest in sports, the weather, cooking, gardening, or ballroom dancing. Identify the attributes, relationships and constraints. What are

the relationship types? You can use the example you thought of yesterday.



e. Describe at least two views of interest on the data. It can be simple and straightforward as long as it includes some entities, attributes and relationships.

The users would be interested in the card entities that include the types, suit, color and value. The users would be interested in this data to try to get some advantage at the table. For instance if the user was at a blackjack table and kept getting high numbers that would be a desirable outcome.

The users would also be interested in the dealer entities and table numbers to see if certain tables had better odds or certain patterns that could be taken into account.

Division of Labor: 33%, 33%, 33%.