many-to-many binary relationships, and ternary relationships (which we consider to in general be many-to-many-to-many.)

Other terms and concepts discussed include cardinality (the *maximum* number of entities that can be involved in a particular relationship), modality (the *minimum* number of entity occurrences that can be involved in a relationship), intersection data (data that describes a many-to-many relationship), and associative entities.

## KEY TERMS

| | | |
|---|---|---|
| Attribute | Entity-relationship (E-R) diagram | One-to-one relationship |
| Associate entity | Entity-relationship (E-R) model | Relationship |
| Binary relationship | Intersection data | Ternary relationship |
| Cardinality | Many-to-many relationship | Unary relationship |
| Data modeling | Modality | Unique identifier |
| Entity | One-to-many relationship | |

## QUESTIONS

1. What is data modeling? Why is it important?
2. What is the Entity-Relationship model?
3. What is a relationship?
4. What are the differences among a unary relationship, a binary relationship, and a ternary relationship?
5. Explain and compare the cardinality of a relationship and the modality of a relationship.
6. Explain the difference between a one-to-one, a one-to-many, and a many-to-many binary relationship.
7. What is intersection data in a many-to-many binary relationship? What does the intersection data describe?
8. Can a many-to-many binary relationship have no intersection data? Explain.
9. Can intersection data be placed in the entity box of one of the two entities in the many-to-many relationship? Explain.

10. What is an associative entity? How does intersection data relate to an associative entity?
11. Describe the three cases of unique identifiers for associative entities.
12. Describe the concept of the unary relationship.
13. Explain how a unary relationship can be described as one-to-one, one-to-many, and many-to-many if only one entity type is involved in the relationship.
14. Describe the ternary relationship concept.
15. Can a ternary relationship have intersection data? Explain.
16. What is a dependent entity? (See the description in the General Hardware example.)

## EXERCISES

1. Draw an entity-relationship diagram that describes the following business environment.

The city of Chicago, IL, wants to maintain information about its extensive system of high schools, including its teachers and their university degrees, its students, administrators, and the subjects that it teaches.

Each school has a unique name, plus an address, telephone number, year built, and size in square feet. Students have a student number, name, home address, home telephone number, current grade, and age. Regarding a student's school assignment, the school system is only interested in keeping track of which school a student *currently* attends. Each school has several administrators, such as the principal and assistant principals. Administrators are identified by an employee number and also have a name, telephone number, and office number.

Teachers are also identified by an employee number and each has a name, age, subject specialty such as English (assume only one per teacher), and the year that they entered the school system. Teachers tend to move periodically from school to school and the school system wants to keep track of the *history* of which schools the teacher has taught in, including the current school. Included will be the year in which the teacher entered the school, and the highest pay rate that the teacher attained at the school. The school system wants to keep track of the universities that each teacher attended, including the degrees earned and the years in which they were earned. The school system wants to record each university's name, address, year founded, and Internet URL (address). Some teachers, as department heads, supervise other teachers. The school system wants to keep track of these supervisory relationships but only for teachers' *current* supervisors.

The school system also wants to keep track of the subjects that it offers (e.g. French I, Algebra III, etc.). Each subject has a unique subject number, a subject name, the grade level in which it is normally taught, and the year in which it was introduced in the school system. The school system wants to keep track of which teacher taught which student which subject, including the year this happened and the grade received.

2. The following entity-relationship diagram describes the business environment of Video Centers of Europe, Ltd., which is a chain of videotape and DVD rental stores. Write a verbal description of how VCE conducts its business, based on this E-R diagram.
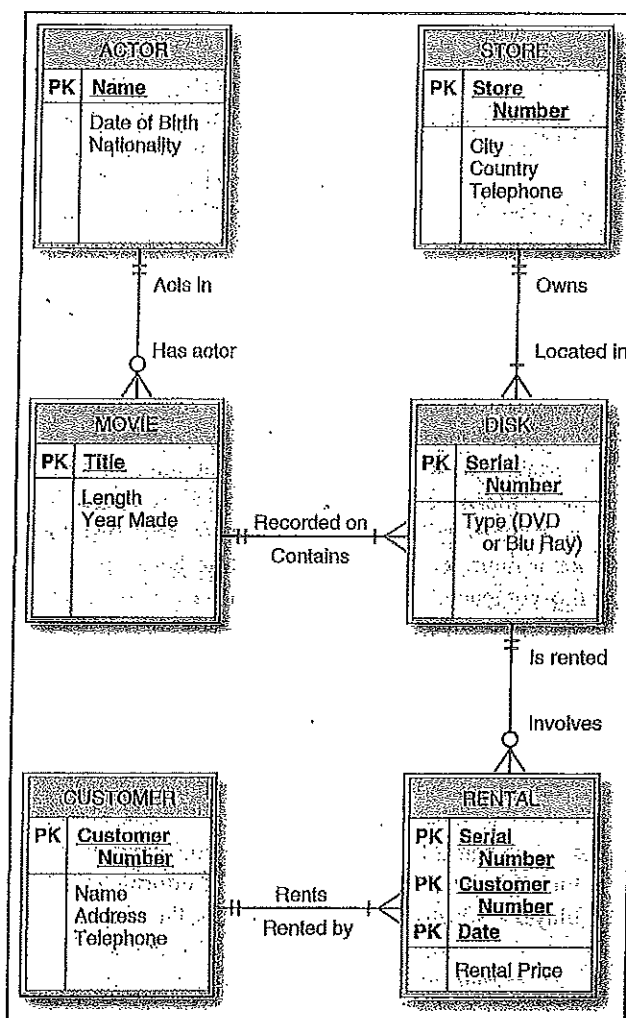


Figure for Exercise 2

# MINICASES

1. Draw an entity-relationship diagram that describes the following business environment.

Happy Cruise Lines has several ships and a variety of cruise itineraries, each involving several ports of call. The company wants to maintain information on the sailors who *currently* work on each of its ships. It also wants to keep track of both its past and future cruises and of the passengers who sailed on the former and are booked on the latter.

Each ship has at least one and, of course, normally many sailors on it. The unique identifier of each ship is its ship number. Other ship attributes include ship name, weight, year built, and passenger capacity. Each sailor has a unique sailor identification number, as well as a name, date of birth, and nationality. Some of the sailors are in supervisory positions, supervising several other sailors. Each sailor reports to just one supervisor. A cruise is identified by a unique cruise serial number. Other cruise descriptors include a sailing date, a return date, and a departure port (which is also the cruise's ending point). Clearly, a cruise involves exactly one ship; over time a ship sails on many cruises, but there

is a requirement to be able to list a new ship that has not yet sailed on any cruises at all. Each cruise stops at at least one and usually several ports of call, each of which is normally host to many cruises, over time. In addition, the company wants to maintain information about ports that it has not yet used in its cruises but may use in the future. A port is identified by its name and the country it is in. Other information about a port includes its population, whether a passport is required for passengers to disembark there, and its current docking fee, which is assumed to be the same for all ships. Passenger information includes a unique passenger number, name, home address, nationality, and date of birth. A cruise typically has many passengers on it (certainly at least one). Hoping for return business, the company assumes that each passenger may have sailed on several of its cruises (and/or may be booked for a future cruise). For a person to be of interest to the company, he or she must have sailed on or be booked on at least one of the company's cruises. The company wants to keep track of how much money each passenger paid (or will pay) for each of their cruises, as well as their satisfaction rating of the cruise, if it has been completed.

2. Draw an entity-relationship diagram that describes the following business environment. The Super Baseball League wants to maintain information about its teams, their coaches, players, and bats. The information about players is historical. For each team, the league wants to keep track of all of the players who have ever played on the team, including the current players. For each player, it wants to know about every team the player ever played for. On the other hand, coach affiliation and bat information is current, only.

The league wants to keep track of each team's team number, which is unique, its name, the city in which it is based, and the name of its manager. Coaches have a name (which is assumed to be unique only within its team) and a telephone number. Coaches have units of work experience that are described by the type of experience and the number of years of that type of experience. Bats are described by their serial numbers (which are unique only within a team) and their manufacturer's name. Players have a player number that is unique across the league, a name, and an age.

A team has at least one and usually several coaches. A coach works for only one team. Each coach has several units of work experience or may have none. Each unit of work experience is associated with the coach to whom it belongs. Each team owns at least one and generally many bats. Currently and historically, each team has and has had many players. To be of interest to the league, a player must have played on at least one and possibly many teams during his career. Further, the league wants to keep track of the number of years that a player has played on a team and the batting average that he compiled on that team.

# TO DO LIST

1. _____
2. _____
3. PG 66 _____  2
4. _____
5. _____
6. _____
7. PG 104 _____  3
8. _____
9. PG -135 _____  4
10. PG 197 _____
11. _____  5
12. _____
13. _____
14. _____
15. _____
16. _____
17. _____
18. _____
19. _____
20. _____

# Albany State University
## Student Academic Calendar
### Spring 2019

| | Full Term | | A Term | | B Term | |
|---|---|---|---|---|---|---|
| Admissions/FA/Readmit/Housing Deadline | November 1, 2018 | | November 1, 2018 | | November 1, 2018 | |
| Advisement/Registration (Current Students)* | November 5, 2018 | November 30, 2018 | November 5, 2018 | November 30, 2018 | November 5, 2018 | November 30, 2018 |
| Fees Due | January 4, 2019 | | January 4, 2019 | | February 13, 2019 | February 14, 2019 |
| Classes Dropped for Non Payment | January 11, 2019 | | January 11, 2019 | | February 15, 2019 | |
| First and Last Day of Semester | January 14, 2019 | May 9, 2019 | January 14, 2019 | March 8, 2019 | February 20, 2019 | |
| Schedule Adjustments (Drop/Add) | January 15, 2019 | January 17, 2019 | January 15, 2019 | January 17, 2019 | March 18, 2019 | May 9, 2019 |
| Enrollment Verifications Due | January 29, 2019 | | January 29, 2019 | | March 19, 2019 | March 20, 2019 |
| Midterms | March 6, 2019 | March 8, 2019 | February 6, 2019 | February 8, 2019 | March 25, 2019 | |
| Midterm Grades Submitted by @ 9:00 AM | March 18, 2019 | | February 11, 2019 | | April 10, 2019 | April 12, 2019 |
| Midterm Grades Available to Students | March 19, 2019 | | February 12, 2019 | | April 15, 2019 | |
| Last day to drop without academic penalty | March 20, 2019 | | February 13, 2019 | | April 16, 2019 | |
| MLK Holiday - Campus Closed | January 21, 2019 | | January 21, 2019 | | April 17, 2019 | |
| Spring Break - No Classes | March 11, 2019 | March 15, 2019 | none | none | none | |
| Founder's Day | April 5, 2019 | | none | | none | |
| Last day of class meetings | May 2, 2019 | | March 5, 2019 | | April 5, 2019 | |
| Student Study Day | May 3, 2019 | | none | | May 2, 2019 | |
| Summer & Fall 2019 Advisement/Registration | TBA | TBA | TBA | TBA | May 3, 2019 | |
| | TBA | TBA | TBA | TBA | TBA | TBA |
| Finals exam Week | May 3, 2019 | May 9, 2019 | March 6, 2019 | March 8, 2019 | TBA | TBA |
| Commencement 10:00 AM | May 11, 2019 | | May 11, 2019 | | May 3, 2019 | May 9, 2019 |
| Grades due including Incompletes due 9:00 AM | May 13, 2019 | | March 18, 2019 | | May 11, 2019 | |
| Grades available to Students | May 17, 2019 | | March 21, 2019 | | May 13, 2019 | |
| | | | | | May 17, 2019 | |

2. Answer the following questions based on the following Super Baseball League data.

| (a) TEAM file. | | | |
|---|---|---|---|
| Team Number | Team Name | City | Manager |
| 137 | Eagles | Orlando | Smith |
| 275 | Cowboys | San Jose | Jones |
| 294 | Statesmen | Springfield | Edwards |
| 368 | Pandas | El Paso | Adams |
| 422 | Sharks | Jackson | Vega |

| (b) PLAYER file. | | | | |
|---|---|---|---|---|
| Player Number | Player Name | Age | Position | Team Number |
| 1209 | Steve Marks | 24 | Catcher | 294 |
| 1254 | Roscoe Gomez | 19 | Pitcher | 422 |
| 1536 | Mark Norton | 32 | First Baseman | 368 |
| 1953 | Alan Randall | 24 | Pitcher | 137 |
| 2753 | John Harbor | 22 | Shortstop | 294 |
| 2843 | John Yancy | 27 | Center Fielder | 137 |
| 3002 | Stuart Clark | 20 | Catcher | 422 |
| 3274 | Lefty Smith | 31 | Third Baseman | 137 |
| 3388 | Kevin Taylor | 25 | Shortstop | 294 |
| 3740 | Juan Vidora | 25 | Catcher | 368 |

a. Regarding the Super Baseball League. Player file shown below.
   i. Describe the file's record type.
   ii. Show a record occurrence.
   iii. Describe the set or range of values that the Player Number field can take.
b. Assume that the records of the Player file are physically stored in the order shown.
   i. Retrieve all of the records of the file physically sequentially.
   ii. Retrieve all of the records of the file logically sequentially based on the Player Name field.
   iii. Retrieve all of the records of the file logically sequentially based on the Player Number field.
   iv. Retrieve all of the records of the file logically sequentially based on the Team Number field.
   v. Perform a direct retrieval of the records with a Player Number field value of 3834.
   vi. Perform a direct retrieval of the records with a Team Number field value of 20.
   vii. Perform a direct retrieval of the records with an Age field value of 24.
c. The value 294 appears as a team number once in the Team file and three times in the Player file. Does this constitute data redundancy? Explain.
d. Merge the Team and Player files based on the common Team Number field (in a manner similar to Figure 3.8 for the General Hardware database). Is the merged file an improvement over the two separate tables in terms of:
   i. Data redundancy? Explain.
   ii. Data integration? Explain.
e. Explain why the Team Number field is in the Player file.
f. Explain why team number 422 appears twice in the Player file.
g. How many files must be accessed to find:
   i. The age of player number 1953?
   ii. The name of the team on which player number 2288 plays?
   iii. The number of the team on which player number 2288 plays?
h. Describe the procedure for finding the name of the city in which player number 3002 is based.
i. What is the mechanism for recording the one-to-many relationship between players and teams in the Super Baseball League database, above?

c. Find the average cost of the motors (see PARTTYPE) supplied by supplier number 3728.
d. List the names of the inspectors who were inspecting refrigerators (see APPLIANCE-TYPE) on April 17, 2011.
e. What was the highest inspection score achieved by a refrigerator on November 3, 2011?
f. Find the total amount of money spent on Quality Appliance products by each buyer from Mexico, Venezuela, and Argentina.

g. Find the total cost of the parts used in each dishwasher manufactured on February 28, 2010. Only include in the results those dishwashers that used at least $200 in parts.
h. List the highest0paid inspectors.
i. List the highest0paid inspectors who were hired in 2009.
j. Among all of the inspectors, list those who earn more than the highest-paid inspector who was hired in 2009.

## MINICASES

1. Consider the following relational database for Happy Cruise Lines. It keeps track of ships, cruises, ports, and passengers. A "cruise" is a particular sailing of a ship on a particular date. For example, the seven-day journey of the ship Pride of Tampa that leaves on June 13, 2011, is a cruise. Note the following facts about this environment.

   - Both ship number and ship name are unique in the SHIP Table.
   - A ship goes on many cruises over time. A cruise is associated with a single ship.
   - A port is identified by the combination of port name and country.
   - As indicated by the VISIT Table, a cruise includes visits to several ports and a port is typically included in several cruises.
   - Both Passenger Number and Social Security Number are unique in the PASSENGER Table. A particular person has a single Passenger Number that is used for all the cruises she takes.
   - The VOYAGE Table indicates that a person can take many cruises and a cruise, of course, has many passengers.

| PORT Table | | | |
| --- | --- | --- | --- |
| PORTNAME | COUNTRY | NUMDOCKS | MANAGER |

| VISIT Table | | | | |
| --- | --- | --- | --- | --- |
| CRUISENUM | PORTNAME | COUNTRY | ARRDATE | DEPDATE |

| PASSENGER Table | | | | |
| --- | --- | --- | --- | --- |
| PASSENGERNUM | PASSENGERNAME | SOCSECNUM | STATE | COUNTRY |

| VOYAGE Table | | | |
| --- | --- | --- | --- |
| PASSENGERNUM | CRUISENUM | ROOMNUM | FARE |

Write SQL SELECT commands to answer the following queries.

a. Find the start and end dates of cruise number 35218.
b. List the names and ship numbers of the ships built by the Ace Shipbuilding Corp. that weigh more than 60,000 tons.
c. List the companies that have built ships for Happy Cruise Lines.
d. Find the total number of docks in all the ports in Canada.
e. Find the average weight of the ships built by the Ace Shipbuilding Corp. that have been launched since 2000.
f. How many ports in Venezuela have at least three docks?

| SHIP Table | | | | |
| --- | --- | --- | --- | --- |
| SHIPNUM | SHIPNAME | BUILDER | LAUNCHDATE | WEIGHT |

| CRUISE Table | | | | |
| --- | --- | --- | --- | --- |
| CRUISENUM | STARTDATE | ENDDATE | DIRECTOR | SHIPNUM |

g. Find the total number of docks in each country. List the results in order from most to least.

h. Find the total number of ports in each country.

i. Find the total number of docks in each country but include only those countries that have at least twelve docks in your answer.

j. Find the name of the ship that operated on (was used on) cruise number 35218.

k. List the names, states and countries of the passengers who sailed on The Spirit of Nashville on cruises that began during July, 2011.

l. Find the names of the company's heaviest ships.

m. Find the names of the company's heaviest ships that began a cruise between July 15, 2011 and July 31, 2011.

2. Consider the following relational database for the Super Baseball League. It keeps track of teams in the league, coaches and players on the teams, work experience of the coaches, bats belonging to each team, and which players have played on which teams. Note the following facts about this environment:

- The database keeps track of the history of all the teams that each player has played on and all the players who have played on each team.
- The database only keeps track of the current team that a coach works for.
- Team number, team name, and player number are each unique attributes across the league.
- Coach name is unique only within a team (and we assume that a team cannot have two coaches of the same name).
- Serial number (for bats) is unique only within a team.
- In the Affiliation table, the Years attribute indicates the number of years that a player played on a team; the batting average is for the years that a player played on a team.

| TEAM Table | | | |
|---|---|---|---|
| TEAMNUM | TEAMNAME | CITY | MANAGER |

| COACH Table | | |
|---|---|---|
| TEAMNUM | COACHNAME | ADDRESS |

| WORK EXPERIENCE Table | | | |
|---|---|---|---|
| TEAMNUM | COACHNAME | EXPERIENCETYPE | YEARSEXPERIENCE |

| BATS Table | | |
|---|---|---|
| TEAMNUM | SERIALNUM | MANUFACTURER |

| PLAYER Table | | |
|---|---|---|
| PLAYERNUM | PLAYERNAME | AGE |

| AFFILIATION Table | | | |
|---|---|---|---|
| PLAYERNUM | TEAMNUM | YEARS | BATTINGAVG |

Write SQL SELECT commands to answer the following queries.

a. Find the names and cities of all of the teams with team numbers greater than 15. List the results alphabetically by team name.

b. List all of the coaches whose last names begin with "D" and who have between 5 and 10 years of experience as college coaches (see YEARSEXPERIENCE and EXPERIENCETYPE).

c. Find the total number of years of experience of Coach Taylor on team number 23.

d. Find the number of different types of experience that Coach Taylor on team number 23 has.

e. Find the total number of years of experience of each coach on team number 23.

f. How many different manufacturers make bats for the league's teams?

g. Assume that team names are unique. Find the names of the players who have played for the Dodgers for at least five years (see YEARS in the AFFILIATION Table.)

h. Assume that team names are unique. Find the total number of years of work experience of each coach on the Dodgers, but include in the result only those coaches who have more than eight years of experience.

i. Find the names of the league's youngest players.

j. Find the names of the league's youngest players whose last names begin with the letter "B".

- Serial Number (for bats) is unique only within a team.
- In the AFFILIATION relation, the Years attribute indicates that number of years that a player played on a team; the Batting Average is for the years that a player played on a team.

| TEAM Relation | | | |
|---|---|---|---|
| Team Number | Team Name | City | Manager |

| COACH Relation | | |
|---|---|---|
| Team Number | Coach Name | Coach Telephone |

| WORK EXPERIENCE Relation | | | |
|---|---|---|---|
| Team Number | Coach Name | Experience Type | Years Of Experience |

| BATS Relation | | |
|---|---|---|
| Team Number | Serial Number | Manufacturer |

| PLAYER Relation | | |
|---|---|---|
| Number Player | Name Player | Age |

| AFFILIATION Relation | | | |
|---|---|---|---|
| Player Number | Team Number | Years | Batting Average |

a. Identify the candidate keys of each relation.
b. Identify the primary key and any alternate keys of each relation.
c. How many foreign keys does each relation have?
d. Identify the foreign keys of each relation.
e. Indicate any instances in which a foreign key serves as part of the primary key of the relation in which it is a foreign key. Why does each of those relations require a multi-attribute primary key?
f. Identify the relations that support many-to-many relationships, the primary keys of those relations, and any intersection data.
g. Assume that we add the following STADIUM relation to the Super Baseball League relational database. Each team has one home stadium, which is what is represented in this relation. Assume that a stadium can serve as the home stadium for only one team. Stadium Name is unique across the league.

| STADIUM Relation | | | |
|---|---|---|---|
| Stadium Name | Year Built | Size | Team Team Number |

What kind of binary relationship exists between the STADIUM relation and the TEAM relation? Could the data from the two relations be combined into one without introducing data redundancy? If so, how?
h. Using the informal relational command language described in this chapter, write commands to:
   i. Retrieve the record for team number 12.
   ii. Retrieve the record for coach Adams on team number 12.
   iii. List the player number and age of every player.
   iv. List the work experience of every coach.
   v. List the work experience of every coach on team number 25.
   vi. Find the age of player number 42459.
   vii. List the serial numbers and manufacturers of all of the Vultures' (the name of a team) bats.
   viii. Find the number of years of college coaching experience that coach Taylor of the Vultures has.

For each of the following tables, first write the table's current normal form (as 1NF, 2NF, or 3NF). Then, reconstruct those tables that are currently in 1 NF or 2 NF as well structured 3 NF tables. Primary key attributes are underlined. Do not assume any functional dependencies other than those shown.

a. State, City, Governor Name, Mayor ID Number, Mayor Name, Mayor Telephone

b. State, City, Mayor Name, Mayor Telephone

c. State, City, Federal Agency, Governor Name, Administrator, Annual Grant

d. State, City, Governor Name, State Flower, Mayor Telephone

e. State, City, City Hall Address, Mayor ID Number, Mayor Name, Mayor Telephone

5. Consider the General Hardware relational database shown in Figure 7.19.

a. Write an SQL command to create the CUSTOMER table.

b. Write an SQL command to create a view of the CUSTOMER table that includes only the Customer Number and HQ City attributes.

c. Write an SQL command to discard the OFFICE table.

d. Assume that Customer Number 8429 is the responsibility of Salesperson Number 758. Write an SQL command to change that responsibility to Salesperson Number 311.

e. Write an SQL command to add a new record to the CUSTOMER table for Customer Number 9442. The Customer Name is Smith Hardware Stores, the responsible salesperson is Salesperson Number 577, and the HQ City is Chicago.

## MINICASES

1. Happy Cruise Lines. Convert the Happy Cruise Lines entity-relationship diagram on the next page into a well structured relational database.

2. Super Baseball League. The Super Baseball League wants to keep track of information about its players, its teams, and the minor league teams (which we will call minor league "clubs" to avoid using the word "team" twice). Minor league clubs are not part of the Super Baseball League but players train in them with the hope of eventually advancing to a team in the Super Baseball League. The intent in this problem is to keep track only of the *current* team on which a player plays in the Super Baseball League. However, the minor league club data must be *historic* and include all of the minor league clubs for which a player has played. Team names, minor league club names, manager names, and stadium names are assumed to be unique, as, of course, is player number.

Design a well structured relational database for this Super Baseball League environment using the data normalization technique. Progress from first to second normal form and then from second to third normal form justifying your design decisions at each step based on the rules of data normalization. The attributes and functional dependencies in this environment are as follows:
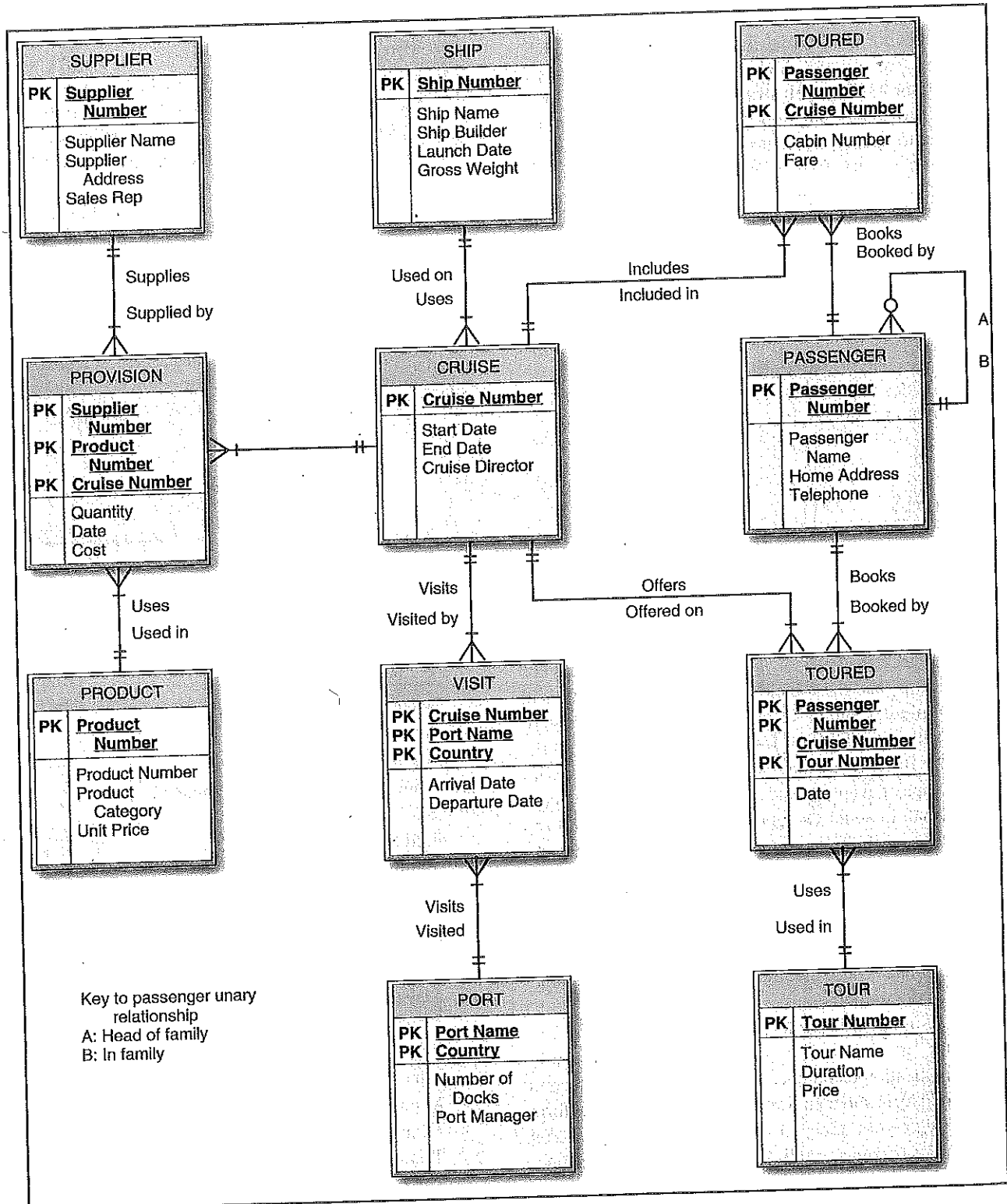
**Attributes**

Player Number

Player Name
Player Age
Team Name
Manager Name
Stadium Name
Minor League Club Name
Minor League Club City
Minor League Club Owner
Minor League Club Year Founded
Start Date
End Date
Batting Average

**Functional Dependencies**

Player Number ———▶ Player Name
Player Number ———▶ Age
Player Number ———▶ Team Name
Player Number ———▶ Manager Name
Player Number ———▶ Stadium Name
Minor League Club Name ———▶ City
Minor League Club Name ———▶ Owner
Minor League Club Name ———▶ Year Founded
Team Name ———▶ Manager Name
Team Name ———▶ Stadium Name
Player Number, Minor League Club Name ———▶ Start Date, End Date, Batting Average

Happy Cruise Lines entity-relationship diagram