1. Exercise 5.2.1 on page 206 of your textbook.
   a.
   
   b. Can you separate the points with at most two per bucket if you use only four grid lines? Either show how or argue that it is not possible.

   We claim that it is NOT POSSIBLE. To prove this, let us first look at the following possibilities for types of grid lines:

   • **0 vertical; 4 horizontal**: This cannot work since no matter how we draw the horizontal grid lines points C, K, and L will be in the same bucket since they all have a speed of 400.

   • **1 vertical, 3 horizontal**: The vertical line will have to be drawn at, or to the right of, 12.7 in order to place points L and \{C,F\} in different buckets. We will return to analyze this positioning further. Note however, that one horizontal line will have to be in the interval (400,450] to put \{C,F\} and \{E,G\} in different buckets; another will have to be in the interval [300,333) to separate \{A,J\} and B and the third will have to be in the interval (266, 300] to put \{A,J\} and I in different buckets.

   Now let’s return to the vertical line. Placing it in the interval (12,12.7] would leave \{B,L,D\} in one bucket. Placing it at, or to the right of 12, however, would leave \{C,F,K\} in one bucket. Consequently there is no place to draw a vertical line without having at least three points in one bucket, eliminating this case.

   • **2 horizontal, 2 vertical**: Note that we will have to place a horizontal line between 300 and 400 to put L and \{A,J\} in different buckets, and we will have to place another between 266 and 300 to place I and \{A,J\} in different buckets. This account for our two horizontal lines. Now we will need to place a vertical line between 6 and 12.7 to place L and \{C,F\} in different buckets. Note that all of the points to the right of this line will be partitioned with at most 2 to a buckets. If we draw the second vertical line to the left of 12.7, however then C, F, D, and K will be in the same bucket, while if we draw the line at, or to the right of, 12.7 then c, f, E, and G will be in the same bucket. Thus, this case is not possible.

   • **1 horizontal, 3 vertical**: Because L, A, J, and I lie on the same vertical line (6) no matter where we draw the horizontal line, at least three point will lie at and above, or at and below this line. Consequently no matter where the vertical lines are drawn we will have at least 3 points in one bucket. Here too, this case is not possible.
- **0 horizontal, 4 vertical**: With no partitioning via horizontal lines, no matter where the vertical lines are draw, \{A, J, L, I\} will lie in the same bucket, eliminating this case as a possibility.

Having ruled out all possible options for placing horizontal and vertical grid lines for partitioning the region into buckets, we have now shown that it is not possible to separate the points with at most two per bucket by using only four grid lines?

c. Suggest a partitioned hash function that will partition these points into four buckets with at most 4 points per bucket.

![Diagram of points and grid lines]

We first note that we can partition the points into four buckets with at most four points per bucket if we draw a horizontal line through 325 and a vertical line through 12.7. Now define

$$h_{\text{speed}}(P) = 0 \text{ if speed < 325 and 1 otherwise}$$
$$h_{\text{size}}(P) = 0 \text{ if size < 12.7 and 1 otherwise.}$$

The $$h(P) = h_{\text{speed}}(P) h_{\text{size}}(P)$$, where the values of $$h_{\text{speed}}$$ and $$h_{\text{size}}$$ are concatenated, yields the desired hash function.

2. Exercise 5.2.3 on page 207 of your textbook.

If we use three bits for a partitioned hash function, let us define $$h_{\text{speed}}$$ and $$h_{\text{size}}$$ as we did in the previous problem, namely

$$h_{\text{speed}}(P) = 0 \text{ if speed < 325 and 1 otherwise}$$
$$h_{\text{size}}(P) = 0 \text{ if size < 12.7 and 1 otherwise.}$$

and now let us define

$$h_{\text{ram}}(P) = 0 \text{ if ram < 64 and 1 otherwise.}$$

If we now define $$h(P) = h_{\text{speed}}(P) h_{\text{ram}}(P) h_{\text{size}}(P)$$ then we get the mappings

<table>
<thead>
<tr>
<th>model - X</th>
<th>h(X)</th>
<th>bucket</th>
<th>contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>000</td>
<td>000</td>
<td>A, H</td>
</tr>
<tr>
<td>B</td>
<td>110</td>
<td>001</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>111</td>
<td>010</td>
<td>I, J</td>
</tr>
<tr>
<td>D</td>
<td>100</td>
<td>011</td>
<td></td>
</tr>
<tr>
<td>E</td>
<td>111</td>
<td>100</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>111</td>
<td>101</td>
<td></td>
</tr>
</tbody>
</table>
3. Exercise 5.2.4 on page 207 of your textbook.

4. Exercise 5.2.5 on page 208 of your etextbook.

Consider the following relational schema for a library (prime attributes are underlined, foreign keys are in italics)

<table>
<thead>
<tr>
<th>Prime Attributes</th>
<th>Foreign Keys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Book(BookID, Title, Publisher)</td>
<td></td>
</tr>
<tr>
<td>Author(AuthorID, Name)</td>
<td></td>
</tr>
<tr>
<td>Branch(BranchID, BranchName, Address)</td>
<td></td>
</tr>
<tr>
<td>BookAuthor(BookID, AuthorID)</td>
<td></td>
</tr>
<tr>
<td>BookCopies(BranchID, BookID, Copies)</td>
<td></td>
</tr>
</tbody>
</table>

5. Express the following queries in SQL:

a. List the title of each book published by Addison-Wesley.

   ```sql
   SELECT Title
   FROM Book
   WHERE Publisher = "Addison-Wesley"
   ```

b. List the title of each book for which GW1 is an author.

   ```sql
   SELECT Title
   FROM Book
   WHERE (BookID IN
            (SELECT BookID
             FROM BookAuthor
             WHERE AuthorID = "GW1"))
   or
   SELECT b.Title
   FROM Book b, BookAuthor a
   WHERE (a.AuthorID = "GW1") AND (b.BookID = a.BookID)
   ```

c. List the title and author of all the books published by Prentice-Hall books that are in the Main library branch.

   ```sql
   SELECT b.Title, a.Name
   FROM Author a, Book b, BookAuthor t, Branch r, BookCopies c
   WHERE (b.Publisher = "Prentice-Hall") AND (r.BranchName = "Main") AND (r.BranchID = c.BranchID) AND (b.Book = c.BookID) AND (c.BookID = t.BookID) AND (a.AuthorID = t.AuthorID);
   ```
d. Give the name and address of each branch that has at least one copy of any book written by G Wiederhold.

```
SELECT BranchName, Address
FROM BRANCH
WHERE BranchID IN
  ( SELECT BranchID
    FROM BOOKCOPIES
    WHERE AuthorID IN
      ( SELECT AuthorID
        FROM AUTHOR
        WHERE Name = "G Wiederhold"
      )
  )
```

e. What are the names of the branches that have at least one copy of all the books written by the author with id GW1?

```
SELECT b.BranchName, b.Address
FROM BRANCH b
WHERE NOT EXISTS
  ( SELECT *
    FROM BOOKAUTHOR a
    WHERE a.AuthorID="GW1"
    AND NOT EXISTS
      ( SELECT *
        FROM BOOKCOPIES c
      )
  )
```

6. Repeat exercise 5. parts a. through d., but use relational algebra instead of SQL. You should use the relational algebra notation given in your book, starting on page 240

a. List the title of each book published by Addison-Wesley.

```
ΠTitle(σPublisher="Addison-Wesley" (BOOK))
```

b. List the title of each book for which GW1 is an author.

```
ΠTitle(BOOK | Π(BookID)(σAuthorID="GW1"(BOOKAUTHOR)))
```

c. List the title and author of all the books published by Prentice-Hall that are in the Main library branch.

```
ΠTitle,Name(AUTHOR | Π(BookID)(σPublisher="Prentice-Hall"(BOOK)) | Π(BookID)(BOOKCOPIES | Π(BranchName="Main" (BRANCH))))
```

d. Give the name and address of each branch that has at least one copy of any book written by G Wiederhold.

```
ΠBranchName, Address (BRANCH | Π(BookID)(BOOKCOPIES | Π(BookID)(BOOKAUTHOR | (σAuthorName="G Wiederhold" (AUTHOR)))))
```