SQL Joins Practice

**Dataset 1 (first 7 problems)**

The dataset consists of four tables: author, book, adaptation, and book\_review.

The first table shows the author data in the following columns:

* id – The author’s unique ID within the database.
* name – The author’s name.
* birth\_year – The year when that author was born.
* death\_year – The year when that author died (the field is empty if they are still alive).

Here are the table’s first few rows:

| **id** | **name** | **birth\_year** | **death\_year** |
| --- | --- | --- | --- |
| 1 | Marcella Cole | 1983 | NULL |
| 2 | Lisa Mullins | 1891 | 1950 |
| 3 | Dennis Stokes | 1935 | 1994 |
| 4 | Randolph Vasquez | 1957 | 2004 |
| 5 | Daniel Branson | 1965 | 1990 |
| … | … | … | … |

The second table, book,  shows details about books. The columns are:

* id – The ID of a given book.
* author\_id – The ID of the author who wrote that book.
* title – The book’s title.
* publish\_year – The year when the book was published.
* publishing\_house – The name of the publishing house that printed the book.
* rating – The average rating for the book.

These are the first five rows:

| **id** | **author\_id** | **title** | **publish\_year** | **publishing\_house** | **rating** |
| --- | --- | --- | --- | --- | --- |
| 1 | NULL | Soulless girl | 2008 | Golden Albatros | 4.3 |
| 2 | NULL | Weak Heart | 1980 | Diarmud Inc. | 3.8 |
| 3 | 4 | Faith Of Light | 1995 | White Cloud Press | 4.3 |
| 4 | NULL | Memory Of Hope | 2000 | Rutis Enterprises | 2.7 |
| 5 | 6 | Warrior Of Wind | 2005 | Maverick | 4.6 |
| … | … | … | … | … | … |

The adaptation table has the following columns:

* book\_id – The ID of the adapted book.
* type – The type of adaptation (e.g. movie, game, play, musical).
* title – The name of this adaptation.
* release\_year – The year when the adaptation was created.
* rating – The average rating for the adaptation.

Here’s a snapshot of the data from this table:

| **book\_id** | **type** | **title** | **release\_year** | **rating** |
| --- | --- | --- | --- | --- |
| 1 | movie | Gone With The Wolves: The Beginning | 2008 | 3 |
| 3 | movie | Companions Of Tomorrow | 2001 | 4.2 |
| 5 | movie | Homeless Warrior | 2008 | 4 |
| 2 | movie | Blacksmith With Silver | 2014 | 4.3 |
| 4 | movie | Patrons And Bearers | 2004 | 3.2 |
| … | … | … | … | … |

The final table is book\_review. It consists of the following columns:

* book\_id - The ID of a reviewed book.
* review - The summary of the review.
* author - The name of the review's author.

Here’s the data:

| **book\_id** | **review** | **author** |
| --- | --- | --- |
| 1 | An incredible book | Sylvia Jones |
| 1 | Great, although it has some flaws | Jessica Parker |
| 2 | Dennis Stokes takes the reader for a ride full of emotions | Thomas Green |
| 3 | Incredible craftsmanship of the author | Martin Freeman |
| 4 | Not the best book by this author | Jude Falth |
| 5 | Claudia Johnson at her best! | Joe Marqiz |
| 6 | I cannot recall more captivating plot | Alexander Durham |

**Problem 1:** Show the name of each author together with the title of the book they wrote and the year in which that book was published.

**Problem 2:** Show the name of each author together with the title of the book they wrote and the year in which that book was published. Show only books published after 2005.

**Problem 3:** For each book, show its title, adaptation title, adaptation year, and publication year.

Include only books with a rating lower than the rating of their corresponding adaptation. Additionally, show only those books for which an adaptation was released within four years of the book’s publication.

Label the title column from the book table book\_title and the title column from the adaptation table adaptation\_title.

**Problem 4:** Show the title of each book together with the title of its adaptation and the date of the release. Show all books, regardless of whether they had adaptations.

**Problem 5:** Show all books with their movie adaptations. Select each book's title, the name of its publishing house, the title of its adaptation, and the type of the adaptation. Keep the books with no adaptations in the result.**:** Show all books with their movie adaptations. Select each book's title, the name of its publishing house, the title of its adaptation, and the type of the adaptation. Keep the books with no adaptations in the result.

**Problem 6:** Join the book\_review and book tables using a RIGHT JOIN. Show the title of the book, the corresponding review, and the name of the review's author. Consider all books, even those that weren't reviewed.

**Problem 7:** Display the title of each book along with the name of its author. Show all books, even those without an author. Show all authors, even those who haven't published a book yet. Use a FULL JOIN.

**Dataset 2 (for the next 3 problems)**

**The first table in the dataset is department. Its columns are:**

* **id – The unique ID of the department.**
* **name – The department name, i.e. where a particular type of product is sold.**

**Here’s the data from the table.**

| **id** | **name** |
| --- | --- |
| **1** | **fruits** |
| **2** | **vegetables** |
| **3** | **seafood** |
| **4** | **deli** |
| **5** | **bakery** |
| **6** | **meat** |
| **7** | **dairy** |

**The second table is product, and it consists of the following columns:**

* **id – The ID of a given product.**
* **name – The product’s name.**
* **department\_id – The ID of the department where the product is located.**
* **shelf\_id – The ID of the shelf of that department where the product is located.**
* **producer\_id – The ID of the company that manufactures this product.**
* **price – The product’s price.**

**Here’s the data snapshot:**

| **id** | **name** | **department\_id** | **shelf\_id** | **producer\_id** | **price** |
| --- | --- | --- | --- | --- | --- |
| **1** | **Apple** | **1** | **1** | **NULL** | **0.5** |
| **2** | **Avocado** | **1** | **1** | **7** | **1** |
| **3** | **Banana** | **1** | **1** | **7** | **0.5** |
| **4** | **Grapefruit** | **NULL** | **1** | **1** | **0.5** |
| **5** | **Grapes** | **1** | **1** | **4** | **2** |
| **…** | **…** | **…** | **…** | **…** | **…** |

**The next table is nutrition\_data. Its columns and data are given below:**

* **product\_id – The ID of a product.**
* **calories – The calorific value of that product.**
* **fat – The amount of fat in that product.**
* **carbohydrate – The amount of carbohydrates in that product.**
* **protein – The amount of protein in that product.**

| **product\_id** | **calories** | **fat** | **carbohydrate** | **protein** |
| --- | --- | --- | --- | --- |
| **1** | **130** | **0** | **5** | **1** |
| **2** | **50** | **4.5** | **3** | **1** |
| **3** | **110** | **0** | **30** | **1** |
| **4** | **60** | **0** | **15** | **1** |
| **NULL** | **90** | **0** | **23** | **0** |
| **…** | **…** | **…** | **…** | **…** |

**The fourth table is named producer. It has the following columns:**

* **id – The ID of a given food producer.**
* **name – The name of the producer.**

**Below is the data from this table:**

| **id** | **name** |
| --- | --- |
| **1** | **BeHealthy** |
| **2** | **HealthyFood Inc.** |
| **3** | **SupremeFoods** |
| **4** | **Foodie** |
| **5** | **Gusto** |
| **6** | **Baker n Sons** |
| **7** | **GoodFoods** |
| **8** | **Tasty n Healthy** |

**The last table in the dataset is sales\_history. It has the following columns:**

* **date – The date of sale.**
* **product\_id – The ID of the product sold.**
* **amount – The amount of that product sold on a particular day.**

**Here’s the data, too:**

| **date** | **product\_id** | **amount** |
| --- | --- | --- |
| **2015-01-14** | **1** | **14** |
| **2015-01-14** | **1** | **13** |
| **2015-01-15** | **2** | **2** |
| **2015-01-16** | **2** | **6** |
| **2015-01-17** | **3** | **8** |
| **…** | **…** | **…** |

**Problem 8::** List all products that have fewer than 150 calories. For each product, show its name (rename the column product) and the name of the department where it can be found (name the column department).

**Problem 9:** For each product, display the:

* Name of the company that produced it (name the column producer\_name).
* Name of the department where the product is located (name it department\_name).
* Product name (name it product\_name).
* Total number of carbohydrates in the product.

Your query should still consider products with no information about producer\_id or department\_id.

**Problem 10:** For each product, show its name, price, producer name, and department name.

Alias the columns as product\_name, product\_price, producer\_name, and department\_name, respectively. Include all the products, even those without a producer or department. Also, include the producers and departments without a product.

**Dataset 3 (for one problem)**

We’ll use the dataset consisting of two tables. The first table is car. Here are its columns:

* id – The car’s ID in the database.
* model – The car’s model.
* brand – The car’s brand.
* original\_price – The original price of that car when new.
* mileage – The car’s total mileage.
* prod\_year – The car’s production year.

The data looks like this:

| **id** | **model** | **brand** | **original\_price** | **mileage** | **prod\_year** |
| --- | --- | --- | --- | --- | --- |
| 1 | Speedster | Teiko | 80,000 | 150,000 | 1999 |
| 2 | Roadmaster | Teiko | 110,000 | 30,000 | 1980 |
| 3 | Sundry | Teiko | 40,000 | 25,000 | 1991 |
| 4 | Furu | Domus | 50,000 | 10,000 | 2002 |
| 5 | Emperor | Domus | 65,000 | 140,000 | 2005 |
| 6 | King | Domus | 200,000 | 6,000 | 1981 |
| 7 | Empress | Domus | 60,000 | 7,600 | 1997 |
| 8 | Fury | Tatsu | 150,000 | 13,000 | 1993 |

The second table is charity\_auction with these columns:

* car\_id – The car’s ID.
* initial\_price – The car’s initial (i.e. starting) price.
* final\_price – The actual price when the car was sold.
* buyer\_id – The ID of the person who bought the car.

Here’s the data:

| **car\_id** | **initial\_price** | **final\_price** | **buyer\_id** |
| --- | --- | --- | --- |
| 1 | 65,000 | NULL | NULL |
| 3 | 35,000 | 50,000 | 1 |
| 5 | 50,000 | 120,000 | 3 |
| 6 | 350,000 | 410,000 | 4 |
| 7 | 65,000 | NULL | NULL |

**Problem 11:** Show the model, brand, and final price of each car sold at the auction. Consider only those sold cars that have more mileage than the car with the id = 4.

**Dataset 4 (for the last problem)**

The dataset for this example consists of only one table: workshop\_workers. It has the following columns.

* id – The worker’s ID.
* name – The worker’s first and last name.
* specialization – The worker's specialization.
* master\_id – The ID of the worker's supervisor.
* experience – The worker's years of experience.
* project\_id – The ID of the project to which the worker is currently assigned.

Here’s the data:

| **id** | **name** | **specialization** | **master\_id** | **experience** | **project\_id** |
| --- | --- | --- | --- | --- | --- |
| 1 | Mathew Conn | woodworking | NULL | 20 | 1 |
| 2 | Kate Brown | woodworking | 1 | 4 | 1 |
| 3 | John Doe | incrusting | 5 | 3 | 1 |
| 4 | John Kowalsky | watchmaking | 7 | 2 | 3 |
| 5 | Suzan Gregowitch | incrusting | NULL | 15 | 4 |

**Problem 12:** Show all workers' names together with the names of their direct supervisors. Rename the columns  apprentice\_name and master\_name, respectively. Consider only workers who have a supervisor (i.e. a master).