1. Given a table of candidates and their skills, you're tasked with finding the candidates best suited for an open .Net MVC developer job. You want to find candidates who are proficient in C#, MVC, and SQL.  
   Write a query to list the candidates who possess all of the required skills for the job. Sort the output by candidate ID in ascending order.

|  |  |
| --- | --- |
| **candidate\_id** | **skill** |
| 001 | C# |
| 001 | MVC |
| 001 | SQL |
| 145 | MVC |
| 145 | PowerBI |
| 145 | SQL |
| 345 | C# |
| 345 | MVC |
| 200 | SQL |

1. Assume you are given the tables below about Instagram **pages** and **page likes**. Write a query to return the page IDs of all the Instagram pages that don't have any likes. The output should be in ascending order.

|  |  |
| --- | --- |
| **page\_id** | **page\_name** |
| 20001 | SQL Solutions |
| 20045 | Brain Exercises |
| 20701 | Tips for Data Analysts |

|  |  |  |
| --- | --- | --- |
| **user\_id** | **page\_id** | **liked\_date** |
| 111 | 20001 | 04/08/2022 00:00:00 |
| 121 | 20045 | 03/12/2022 00:00:00 |
| 156 | 20001 | 07/25/2022 00:00:00 |

1. Write a query to get the number of companies that have posted duplicate job listings. **Clarification:** Duplicate job listings refer to two jobs at the same company with the same title and description.

|  |  |  |  |
| --- | --- | --- | --- |
| **job\_id** | **company\_id** | **title** | **description** |
| 248 | 827 | Business Analyst | Business analyst evaluates past and current business data with the primary goal of improving decision-making processes within organizations. |
| 149 | 845 | Business Analyst | Business analyst evaluates past and current business data with the primary goal of improving decision-making processes within organizations. |
| 945 | 345 | Data Analyst | Data analyst reviews data to identify key insights into a business's customers and ways the data can be used to solve problems. |
| 164 | 345 | Data Analyst | Data analyst reviews data to identify key insights into a business's customers and ways the data can be used to solve problems. |
| 172 | 244 | Data Engineer | Data engineer works in a variety of settings to build systems that collect, manage, and convert raw data into usable information for data scientists and business analysts to interpret. |

1. Assume you are given the table below on Uber transactions made by users. Write a query to obtain the third transaction of every user. Output the user id, spend and transaction date.

|  |  |  |
| --- | --- | --- |
| **user\_id** | **spend** | **transaction\_date** |
| 111 | 100.50 | 01/08/2022 12:00:00 |
| 111 | 55.00 | 01/10/2022 12:00:00 |
| 121 | 36.00 | 01/18/2022 12:00:00 |
| 145 | 24.99 | 01/26/2022 12:00:00 |
| 111 | 89.60 | 02/05/2022 12:00:00 |

1. Assume there are three Spotify tables containing information about the artists, songs, and music charts. Write a query to determine the top 5 artists whose songs appear in the Top 10 of the global\_song\_rank table the highest number of times. From now on, we'll refer to this ranking number as "song appearances".  
   Output the top 1 artist names with their song appearances ranking (not the number of song appearances, but the rank of who has the most appearances). The order of the rank should take precedence.  
   For example, Ed Sheeran's songs appeared 5 times in Top 10 list of the global song rank table; this is the highest number of appearances, so he is ranked 1st. Bad Bunny's songs appeared in the list 4, so he comes in at a close 2nd

|  |  |
| --- | --- |
| **artist\_id** | **artist\_name** |
| 101 | Ed Sheeran |
| 120 | Drake |

|  |  |
| --- | --- |
| **song\_id** | **artist\_id** |
| 45202 | 101 |
| 19960 | 120 |

|  |  |  |
| --- | --- | --- |
| **day** | **song\_id** | **rank** |
| **1** | **45202** | **5** |
| **3** | **45202** | **2** |
| **1** | **19960** | **3** |
| **9** | **19960** | **15** |

1. Assume you are given the table on Walmart user transactions. Based on a user's most recent transaction date, write a query to obtain the users and the number of products bought.  
   Output the user's most recent transaction date, user ID and the number of products sorted by the transaction date in chronological order.

|  |  |  |  |
| --- | --- | --- | --- |
| **product\_id** | **user\_id** | **spend** | **transaction\_date** |
| 3673 | 123 | 68.90 | 07/08/2022 12:00:00 |
| 9623 | 123 | 274.10 | 07/08/2022 12:00:00 |
| 1467 | 115 | 19.90 | 07/08/2022 12:00:00 |
| 2513 | 159 | 25.00 | 07/08/2022 12:00:00 |
| 1452 | 159 | 74.50 | 07/10/2022 12:00:00 |

1. You’re a consultant for a major pizza chain that will be running a promotion where all 3-topping pizzas will be sold for a fixed price, and are trying to understand the costs involved.Given a list of pizza toppings, consider all the possible 3-topping pizzas, and print out the total cost of those 3 toppings. Sort the results with the highest total cost on the top followed by pizza toppings in ascending order.

Break ties by listing the ingredients in alphabetical order, starting from the first ingredient, followed by the second and third.

|  |  |
| --- | --- |
| **topping\_name** | **ingredient\_cost** |
| Pepperoni | 0.50 |
| Sausage | 0.70 |
| Chicken | 0.55 |
| Extra Cheese | 0.40 |

1. Assume you have the table below containing information on Facebook user actions. Write a query to obtain the active user retention in May 2022. Output the month (in numerical format 1, 2, 3) and the number of monthly active users (MAUs). Hint: An active user is a user who has user action ("sign-in", "like", or "comment") in the current month and last month.

|  |  |  |  |
| --- | --- | --- | --- |
| **user\_id** | **event\_id** | **event\_type** | **event\_date** |
| 445 | 7765 | sign-in | 05/31/2022 12:00:00 |
| 742 | 6458 | sign-in | 06/03/2022 12:00:00 |
| 445 | 3634 | like | 06/05/2022 12:00:00 |
| 742 | 1374 | comment | 06/05/2022 12:00:00 |
| 648 | 3124 | like | 06/18/2022 12:00:00 |