Q1. Cities With Completed Trades

Problem Statement:

Write a query to list the **top three** cities that have the most completed trade orders along with the total number of orders.

- Return the result table ordered by `total_orders` in **descending order.**

Sample Input:

Table: `trades`

order_id	user_id	price	quantity	status	timestamp
100101	111	9.8	10	Cancelled	08/17/2022 12:00:00
100102	111	10	10	Completed	08/17/2022 12:00:00
100259	148	5.1	35	Completed	08/25/2022 12:00:00
100264	148	4.8	40	Completed	08/26/2022 12:00:00
100305	300	10	15	Completed	09/05/2022 12:00:00
100400	178	9.9	15	Completed	09/09/2022 12:00:00
100565	265	25.6	5	Completed	12/19/2022 12:00:00

Table: `users`

user_id	city	email	signup_date
111	San Francisco	rrok10@gmail.com	08/03/2021 12:00:00
148	Boston	sailor9820@gmail.com	08/20/2021 12:00:00
178	San Francisco	harrypotterfan182@gmail.com	01/05/2022 12:00:00
265	Denver	shadower_@hotmail.com	02/26/2022 12:00:00
300	San Francisco	houstoncowboy1122@hotmail.com	06/30/2022 12:00:00

Sample output:

city	total_orders
San Francisco	3
Boston	2
Denver	1

Explanation:

San Francisco has first place with 3 orders, Boston has second place with 2 orders, and Denver has third place with 1 order.

Q2. Running Total for Different Genders

Problem Statement:

Write a query to find the running total score for each gender on each day.

- Return the result table ordered by `gender` and `day` in **ascending** order.

Sample Input:

Table: scores

player_name	gender	day	score_points
Aron	F	2020-01-01	17
Alice	F	2020-01-07	23
Bajrang	M	2020-01-07	7
Khali	M	2019-12-25	11
Slaman	M	2019-12-30	13
Joe	M	2019-12-31	3
Jose	M	2019-12-18	2
Priya	F	2019-12-31	23
Priyanka	F	2019-12-30	17

Sample output:

gender	day	total
F	2019-12-30	17
F	2019-12-31	40
F	2020-01-01	57
F	2020-01-07	80
M	2019-12-18	2
M	2019-12-25	13
M	2019-12-30	26
M	2019-12-31	29
M	2020-01-07	36

Explanation:

- **For the female team:**
  - The first day is 2019-12-30, Priyanka scored 17 points and the total score for the team is 17.
  - The second day is 2019-12-31, Priya scored 23 points and the total score for the team is 40.
  - The third day is 2020-01-01, Aron scored 17 points and the total score for the team is 57.
  - The fourth day is 2020-01-07, Alice scored 23 points and the total score for the team is 80.
- **For the male team:**
  - The first day is 2019-12-18, Jose scored 2 points and the total score for the team is 2.
  - The second day is 2019-12-25, Khali scored 11 points and the total score for the team is 13.
  - The third day is 2019-12-30, Slaman scored 13 points and the total score for the team is 26.
  - The fourth day is 2019-12-31, Joe scored 3 points and the total score for the team is 29.
  - The fifth day is 2020-01-07, Bajrang scored 7 points and the total score for the team is 36.

Q3. Countries You Can Safely Invest In

Problem Statement:

A telecommunications company wants to invest in new countries. The company intends to invest in countries where the **average call duration** of the calls in this country is **strictly greater** than the global average call duration.

Write a query to find the countries where this company can invest.

- Return the result table ordered by `country` column in ascending order.

Sample Input:

Table: person

id	name	phone_number
3	Jonathan	051-1234567
12	Elvis	051-7654321
1	Moncef	212-1234567
2	Maroua	212-6523651
7	Meir	972-1234567
9	Rachel	972-0011100

Table: country

name	country_code
Peru	51
Israel	972
Morocco	212
Germany	49
Ethiopia	251

Table: calls

caller_id	callee_id	duration
1	9	33
2	9	4
1	2	59
3	12	102
3	12	330
12	3	5
7	9	13
7	1	3
9	7	1
1	7	7

Sample output:

country
Peru

Explanation:

- The average call duration for Peru is (102 + 102 + 330 + 330 + 5 + 5) / 6 = 145.666667
- The average call duration for Israel is (33 + 4 + 13 + 13 + 3 + 1 + 1 + 7) / 8 = 9.37500
- The average call duration for Morocco is (33 + 4 + 59 + 59 + 3 + 7) / 6 = 27.5000
- Global call duration average = (2 * (33 + 4 + 59 + 102 + 330 + 5 + 13 + 3 + 1 + 7)) / 20 = 55.70000
- Since Peru is the only country where the average call duration is greater than the global average, it is the only recommended country.

Q4. Same salaries

Context:

TechCorp Innovations is dedicated to delivering cutting-edge technology solutions and promoting a collaborative work environment to drive innovation and growth. The company is restructuring its workforce to enhance team collaboration and productivity. As part of this initiative, the company plans to group employees into teams based on their salaries. This strategy will help ensure that employees with similar compensation levels work together, fostering a sense of unity and shared goals.

You are an analyst at TechCorp Innovations. Your manager has requested a report on the formation of teams based on employee salaries. This report will assist in the reorganization efforts and help HR manage teams effectively.

Problem Description:

A company wants to divide the employees into teams such that all the members of each team have the same salary. The teams should follow these criteria:

- Each team should consist of **at least two** employees.
- All the employees on a team should have the **same** salary.
- All the employees with the **same salary** should be assigned to the **same team**.
- If an employee's salary is **unique**, we **do not assign** this employee to any team.
- A team's ID is assigned based on the rank of the team's salary relative to the other teams' salaries, where the team with the **lowest salary** has team_id = **1**.
- Note that the salaries for employees not on a team are not included in this ranking.

Write a query to get the **team_id** of each employee that is in a team.

Result:

- Return the result table ordered by **team_id** in ascending order. In case of a tie, order it by **employee_id** in ascending order.

Sample Input:

Table: employees

employee_id	name	salary
1	Andrew	5000
2	Erin	5000
3	Stanley	4700
4	Jim	8000
5	Oscar	8000

Sample Output:

employee_id	name	salary	team_id
1	Andrew	5000	1
2	Erin	5000	1
4	Jim	8000	2
5	Oscar	8000	2

Sample Explanation:

- Andrew (employee_id=1) and Erin (employee_id=2) are on the same team because they have the same salary of 5000.
- Jim (employee_id=4) and Oscar (employee_id=5) are on the same team because they have the same salary of 8000.
- Stanley (employee_id=3) is not included in any team because their salary of 4700 is unique (i.e. no other employee has the same salary).
- The team IDs are assigned as follows (based on salary ranking, lowest first):
  - team_id=1: Andrew and Erin, a salary of 5000
  - team_id=2: Jim and Oscar, a salary of 8000
- Stanley's salary of 4700 is not included in the ranking because he is not on a team.

Q5. 5th highest – Use hr dataset

Context:

GlobalTech Solutions is a leading technology company specializing in software development and IT services. The company is dedicated to delivering cutting-edge technology solutions and maintaining an equitable work environment for its employees. As part of their employee management system, they need to analyze salary data to ensure competitive compensation across different job categories.

You are a data analyst at GlobalTech Solutions. Your manager has asked you to identify employees as per the problem statement within each job category. This analysis will help the company maintain a fair and competitive salary structure.

Problem Statement:

Write a query to display the details of the employees who have the **5th highest** salary in each job category.

Result:

- Return the columns '**employee_id**', '**first_name**', and '**job_id**'.
- Return the result ordered by **employee_id** in ascending order.

Sample Input: (Use hr dataset)

Table: employees

employee_id	first_name	last_name	email	phone_number	hire_date	job_id	salary	commission_pct	manager_id	department_id
109	Daniel	Faviet	DFAVIET	515.124.4169	1994-08-16	FI_ACCOUNT	9000	NULL	108	100
110	John	Chen	JCHEN	515.124.4269	1997-09-28	FI_ACCOUNT	8200	NULL	108	100
111	Ismael	Sciarra	ISCIARRA	515.124.4369	1997-09-30	FI_ACCOUNT	7700	NULL	108	100
112	Jose Manuel	Urman	JMURMAN	515.124.4469	1998-03-07	FI_ACCOUNT	7800	NULL	108	100
113	Luis	Popp	LPOPP	515.124.4567	1999-12-07	FI_ACCOUNT	6900	NULL	108	100
115	Alexander	Khoo	AKHOO	515.127.4562	1995-05-18	PU_CLERK	3100	NULL	114	30
116	Shelli	Baida	SBAIDA	515.127.4563	1997-12-24	PU_CLERK	2900	NULL	114	30
117	Sigal	Tobias	STOBIAS	515.127.4564	1997-07-24	PU_CLERK	2800	NULL	114	30
118	Guy	Himuro	GHIMURO	515.127.4565	1998-11-15	PU_CLERK	2600	NULL	114	30
119	Karen	Colmenares	KCOLMENA	515.127.4566	1999-08-10	PU_CLERK	2500	NULL	114	30

Sample Output:

employee_id	first_name	job_id
113	Luis	FI_ACCOUNT
119	Karen	PU_CLERK

Q6. Winning streak

Problem Description:

The winning streak of a player is calculated as the number of consecutive wins uninterrupted by draws or losses.

Write a query to count the **longest winning streak** for **each** player and save the new column as 'longest_streak'.

- Return the **player_id** and **longest_streak.**
- Return the result ordered by **player_id** in ascending order.

Sample Input:

Table: matches

player_id	match_day	result
1	2022/07/13	Win
1	2022/07/14	Win
1	2022/07/16	Win
1	2022/07/18	Draw
1	2022/07/20	Win
2	2022/07/18	Lose
2	2022/07/19	Lose
3	2022/07/18	Win
3	2022/07/21	Win
3	2022/07/22	Lose
3	2022/07/23	Lose

Sample Output:

player_id	longest_streak
1	3
2	0
3	2

Sample Explanation:

- Player 1:
    - From 2022/07/13 to 2022/07/16, player 1 won 3 consecutive matches.
    - On 2022/07/18, player 1 had a draw.
    - On 2022/07/20, player 1 won a match.
    - The longest winning streak was 3 matches.
- Player 2:
    - From 2022/07/18 to 2022/07/19, player 2 lost 2 consecutive matches.
    - The longest winning streak was 0 matches.
- Player 3:
    - From 2022/07/18 to 2022/07/21, player 3 won 2 matches.
    - From 2022/07/22 to 2022/07/23, player 3 lost 2 matches.
    - The longest winning streak was 2 matches.