Q1. Department wise salary analysis – 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 its employee management system, to ensure fair and transparent employee evaluations, the company wants to identify and reward top-performing employees within each department.

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

Problem Statement:

Write a query to calculate

- the **row number** and save it as 'emp_row_no',
- **rank** and save as 'emp_rank', and
- the **dense rank** of employees as 'emp_dense_rank'

based on the **salary** column in **descending order within each department** using the employees table.

Result:

- Return the columns '**full_name**' (first_name and last_name separated by space), '**department_id**', '**salary**', '**emp_row_no**', '**emp_rank**', and '**emp_dense_rank**'.
- Return the result ordered by **department_id** in ascending order and **salary** in descending order.

Sample Input:

Table: employees

employee_id	first_name	last_name	email	phone_number	hire_date	job_id	salary	commission_pct	manager_id	department_id
125	Julia	Nayer	JNAYER	650.124.1214	1997-07-16	ST_CLERK	3200	NULL	120	50
129	Laura	Bissot	LBISSOT	650.124.5234	1997-08-20	ST_CLERK	3300	NULL	121	50
133	Jason	Mallin	JMALLIN	650.127.1934	1996-06-14	ST_CLERK	3300	NULL	122	50
134	Michael	Rogers	MROGERS	650.127.1834	1998-08-26	ST_CLERK	2900	NULL	122	50
137	Renske	Ladwig	RLADWIG	650.121.1234	1995-07-14	ST_CLERK	3600	NULL	123	50
141	Trenna	Rajs	TRAJS	650.121.8009	1995-10-17	ST_CLERK	3500	NULL	124	50
186	Julia	Dellinger	JDELLING	650.509.3876	1998-06-24	SH_CLERK	3400	NULL	121	50
188	Kelly	Chung	KCHUNG	650.505.1876	1997-06-14	SH_CLERK	3800	NULL	122	50
189	Jennifer	Dilly	JDILLY	650.505.2876	1997-08-13	SH_CLERK	3600	NULL	122	50
200	Jennifer	Whalen	JWHALEN	515.123.4444	1987-09-17	AD_ASST	4400	NULL	101	10
201	Michael	Hartstein	MHARTSTE	515.123.5555	1996-02-17	MK_MAN	13000	NULL	100	20
202	Pat	Fay	PFAY	603.123.6666	1997-08-17	MK_REP	6000	NULL	201	20

Sample Output:

full_name	department_id	salary	emp_row_no	emp_rank	emp_dense_rank
Jennifer Whalen	10	4400	1	1	1
Michael Hartstein	20	13000	1	1	1
Pat Fay	20	6000	2	2	2
Kelly Chung	50	3800	1	1	1
Renske Ladwig	50	3600	2	2	2
Jennifer Dilly	50	3600	3	2	2
Trenna Rajs	50	3500	4	4	3
Julia Dellinger	50	3400	5	5	4
Laura Bissot	50	3300	6	6	5
Jason Mallin	50	3300	7	6	5
Julia Nayer	50	3200	8	8	6
Michael Rogers	50	2900	9	9	7

Q2. Quartile – Use HR Dataset

Context:

Retail Prime, a leading e-commerce platform, strives to ensure a fair and competitive compensation structure for its employees. Analyzing salary distribution and identifying quartiles helps the company maintain equitable pay scales and make informed decisions about salary adjustments and promotions.

As a data analyst at Retail Prime, you are tasked with generating a report that categorizes each employee's salary into quartiles. This will assist in understanding the salary distribution across different departments and roles within the company.

Problem Statement:

Find the **quartile** of each record based on the **salary** of the employee and save it as 'Quartile'.

Result:

- Return the columns '**employee_id**', '**first_name**', '**department_id**', '**job_id**', '**salary**', '**Quartile**'.
- **Quartile** refers to any of the **4** equal bins into which a data set can be divided, based on the values of a particular variable.
- Return the output ordered by the **Quartile**, **salary**, and **employee_id** columns in ascending order.

Sample Input:
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	department_id	job_id	salary	Quartile
119	Karen	30	PU_CLERK	2500	1
118	Guy	30	PU_CLERK	2600	1
117	Sigal	30	PU_CLERK	2800	1
116	Shelli	30	PU_CLERK	2900	2
115	Alexander	30	PU_CLERK	3100	2
113	Luis	100	FI_ACCOUNT	6900	2
111	Ismael	100	FI_ACCOUNT	7700	3
112	Jose Manuel	100	FI_ACCOUNT	7800	3
110	John	100	FI_ACCOUNT	8200	4
109	Daniel	100	FI_ACCOUNT	9000	4

Sample Explanation:
In this sample output, each employee's salary is categorized into quartiles based on the distribution of all salaries. The salaries are divided into four equal bins:

- The 1st quartile (Q1) includes the lowest 25% of salaries: 2500, 2600, 2800.
- The 2nd quartile (Q2) includes the next 25% of salaries: 2900, 3100, 6900.
- The 3rd quartile (Q3) includes the next 25% of salaries: 7700, 7800.
- The 4th quartile (Q4) includes the highest 25% of salaries: 8200, 9000

Each employee's salary is placed into one of these quartiles.
Here, Karen, Guy, and Sigal fall into the 1st quartile, while John and Daniel fall into the 4th quartile.
The output lists the employee_id, first_name, department_id, job_id, salary, and Quartile for each employee, ordered by Quartile, salary, and employee_id in ascending order.

Q3. Human traffic

Context:

UrbanMall is dedicated to understanding customer traffic patterns to enhance mall operations and optimize resource allocation. Identifying peak traffic periods where multiple consecutive visits had high attendance is crucial for making strategic decisions on staffing and promotions.

As a data analyst at UrbanMall, your task is to identify and report instances where there were three or more consecutive visits with a significant number of people, specifically 100 or more, during each visit. This analysis will help in recognizing patterns of high traffic and planning accordingly.

Problem Statement:

- Each row in the table contains the visit_date and visit_id to the mall with the number of people during the visit. No two rows will have the same visit_date
- Write a query to find and display records where there are **three or more** consecutive visits with each having at least 100 people.

Result:

- Return the columns **visit_id**, **visit_date**, and **people**.
- Return the result ordered by **visit_date** in ascending order.

Sample Input:

Table: mall

id	visit_date	people
1	2022/07/13	50
2	2022/07/14	190
3	2022/07/15	20
4	2022/07/16	300
5	2022/07/18	450
6	2022/07/19	600
7	2022/07/20	110
8	2022/07/21	220

Sample Output:

id	visit_date	people
4	2022/07/16	300
5	2022/07/18	450
6	2022/07/19	600
7	2022/07/20	110
8	2022/07/21	220

Sample Explanation:

- The five rows with ids 4, 5, 6, 7, and 8 have consecutive ids, and each of them has ≥ 100 people attended. Note that row 5 was included even though the visit_date was not the next day after row 4 (we check only with consecutive IDs which may and may not be consecutive dates).
- The rows with ids 1 and 3 are not included because we need ≥ 100 people.
- The row with id 2 was not included because we need three or more rows with consecutive IDs.

Q4. Salaries after taxes

Context:

TechWave Solutions is a multinational corporation specializing in innovative technology solutions and software development. The company aims to ensure competitive and fair employee compensation while accounting for tax regulations that vary based on salary brackets. By understanding the effective post-tax salaries of employees, TechWave Solutions can better plan its financial strategies and employee benefits.You are a finance analyst at TechWave Solutions. Your manager has requested a report on the post-tax salaries of employees, considering the different tax rates applicable based on the maximum salary within each company. This information will help in budgeting and ensuring competitive compensation packages.

Problem Description:
Write a query to find the **salaries** of the employees **after applying taxes**. Round the salary to the nearest integer. The tax rate is calculated for each company based on the following criteria:

- 0% If the max salary of **any** employee in the company is **less than** $1000.
- 24% If the max salary of **any** employee in the company is in the **range** [1000, 10000] inclusive.
- 49% If the max salary of **any** employee in the company is **greater than** $10000.

Result:

1. **The salary after taxes** = salary - salary x (taxes percentage / 100).
2. Return the result ordered by **company_id**, and **employee_id** in ascending order.

Sample Input:

Table: salaries

company_id	employee_id	employee_name	salary
1	1	Andrew	13725
1	2	Erin	5366
1	3	Stanley	8800
1	4	Jim	14609
1	5	Oscar	14551
2	1	Kevin	100
2	2	Pam	800
2	3	Ryan	600
2	4	Kelly	750

Sample Output:

company_id	employee_id	employee_name	salary
1	1	Andrew	7000
1	2	Erin	2737
1	3	Stanley	4488
1	4	Jim	7451
1	5	Oscar	7421
2	1	Kevin	100
2	2	Pam	800
2	3	Ryan	600
2	4	Kelly	750

Sample Explanation:

- For company 1, the **max salary** is 14609. Employees in company 1 have **taxes = 49%**
- For company 2, the **max salary** is 800. Employees in company 2 have **taxes = 0%**

Q5. Account Balance

Context:

At FinTech Solutions, an innovative financial services company, tracking the balance of user accounts after each transaction is essential for maintaining accurate financial records and ensuring the integrity of transactions. Understanding the balance changes over time helps in detecting anomalies and providing users with transparent account activity.

As a data analyst at FinTech Solutions, your manager has tasked you with generating a report showing each user's balance after every transaction. This will help the finance team monitor account activity and ensure compliance with financial regulations.

Problem Statement:

Write a query to report the balance of each user after each transaction. You may assume that the balance of each account before any transaction is 0 and that the balance will never be below 0 at any moment.

Result:

- Return the columns '**account_id**', '**day',** and '**balance**'.
- Return the result table in **ascending order** by **account_id**, then by **day** in case of a tie.

Sample Input:

Table: transactions

account_id	day	type	amount
1	2021-11-07	Deposit	2000
1	2021-11-09	Withdraw	1000
1	2021-11-11	Deposit	3000
2	2021-12-07	Deposit	7000
2	2021-12-12	Withdraw	7000

Sample output:

account_id	day	balance
1	2021-11-07	2000
1	2021-11-09	1000
1	2021-11-11	4000
2	2021-12-07	7000
2	2021-12-12	0

Explanation:

- Account 1:
  - Initial balance is 0.
  - 2021-11-07 --> deposit 2000. Balance is 0 + 2000 = 2000.
  - 2021-11-09 --> withdraw 1000. Balance is 2000 - 1000 = 1000.
  - 2021-11-11 --> deposit 3000. Balance is 1000 + 3000 = 4000.
- Account 2:
  - Initial balance is 0.
  - 2021-12-07 --> deposit 7000. Balance is 0 + 7000 = 7000.
  - 2021-12-12 --> withdraw 7000. Balance is 7000 - 7000 = 0.