SQL PROJECT FOR TECH SOLUTIONS

datalab

Tech Solutions Inc. is a leading technology company specializing in software development and IT consulting services. The company prides itself on delivering innovative solutions to clients across various industries. With a dedicated team of skilled professionals, TechSolutions has earned a reputation for excellence in the tech industry.

Tech Solutions Inc. has been experiencing a decline in customer satisfaction ratings over the past few months. Customer feedback surveys and support tickets indicate an increase in dissatisfaction among clients. The company is concerned about this trend as it directly impacts customer retention, reputation, and overall business growth.

You are working with the customer support team to provide data to managers to help the company take proactive measures to address these concerns effectively.

Data

The following schema diagram shows the tables available.

Task 1

Before you can start any analysis, you need to confirm that the data is accurate and reflects what you expect to see.

It is known that there are some issues with the `support` table, and the data team have provided the following data description.

Write a query to return data matching this description. You must match all column names and description criteria.

Column Name	Criteria
id	Discrete. The unique identifier of the support ticket. Missing values are not possible due to the database structure.
customer_id	Discrete. The unique identifier of the customer. Missing values should be replaced with 0.
category	Nominal. The gategory of the support request, can be one of Feedback, Billing Enquiry, Bug, Installation Problem, Other. Missing values should be replaced with Other.
status	Nominal. The current status of the support ticket, one of Open, In Progress or Resolved. Missing values should be replaced with 'Resolved'.
creation_date	Discrete. The date the ticket was created. Can be any date in 2023. Missing values should be replaced with 2023-01-01.
response_time	Discrete. The number of days taken to respond to the support ticket. Missing values should be replaced with 0.
resolution_time	Continuos. The number of hours taken to resolve the support ticket, rounded to 2 decimal places. Missing values should be replaced with 0.

Unknown integration DataFrame as `clean_data`

```sql
-- Write your query for task 1 in this cell
SELECT
    id,
    COALESCE(customer_id, 0) AS customer_id,
    COALESCE(category, 'Other') AS category,
    CASE
        WHEN status IS NULL THEN 'Resolved'
        WHEN status = '-' THEN 'Resolved'
        ELSE status
    END AS status,
    COALESCE(creation_date, '2023-01-01'::DATE) AS creation_date,
    COALESCE(response_time, 0) AS response_time,
COALESCE(ROUND(NULLIF(regexp_replace(resolution_time, '[^0-9.]', '', 'g'), '')::NUMERIC, 2), 0) AS resolution_time
FROM public.support;
```

i..	i..	custome...	category	status	creation_date	res
0	1	1062	Installation Problem	In Progress	2023-01-26T00:00:00.000	
1	2	892	Billing enquiry	Open	2023-06-18T00:00:00.000	
2	3	433	Feedback	Open	2023-08-17T00:00:00.000	
3	6	764	Billing enquiry	Open	2023-01-16T00:00:00.000	
4	7	1144	Billing enquiry	Open	2023-06-01T00:00:00.000	
5	8	288	Feedback	Open	2023-01-22T00:00:00.000	
6	9	1495	Bug	In Progress	2023-02-05T00:00:00.000	
7	10	1090	Bug	In Progress	2023-05-09T00:00:00.000	
8	11	1397	Feedback	In Progress	2023-09-17T00:00:00.000	
9	12	54	Feedback	Open	2023-09-13T00:00:00.000	
10	14	1207	Bug	Resolved	2023-09-04T00:00:00.000	
11	15	1256	Bug	In Progress	2023-11-20T00:00:00.000	
12	16	146	Feedback	Open	2023-09-08T00:00:00.000	
13	17	455	Installation Problem	In Progress	2023-11-07T00:00:00.000	
14	19	950	Bug	In Progress	2023-11-21T00:00:00.000	
15	21	215	Feedback	Open	2023-05-22T00:00:00.000	

Rows: 1,987 Expand

Task 2

It is suspected that the response time to tickets is a big factor in unhappiness.

Calculate the minimum and maximum response time for each category of support ticket.

Your output should include the columns `category`, `min_response` and `max_response`.

Values should be rounded to two decimal places where appropriate.

Unknown integration DataFrame as `min_max_response`

```sql
-- Write your query for task 2 in this cell
SELECT category,
       ROUND(MIN(response_time),2) AS min_response,
       ROUND(MAX(response_time),2) AS max_response
FROM public.support
GROUP BY category;
```

	category	min_...	max...
0	Other	1	5
1	Bug	1	13
2	Feedback	1	2
3	Billing enquiry	2	8
4	Installation Problem	5	17

Rows: 5 Expand

Task 3

The support team want to know more about the `rating` provided by customers who reported `Bugs` or `Installation Problem`s.

Write a query to return the `rating` from the survey, the `customer_id`, `category` and `response_time` of the support ticket, for the customers & categories of interest.

Use the original support table, not the output of task 1.

Unknown integration DataFrame as t

```
SELECT s2.rating AS rating,
       s1.customer_id AS customer_id,
       s1.category AS category,
       s1.response_time AS response_time
FROM public.support s1
JOIN public.survey s2
ON s2.id = s1.id OR s2.customer_id=s1.customer_id
WHERE s1.category IN ('Bugs', 'Installation Problem');
```

	rating	cus...	category	respon...
0	3	1062	Installation Problem	6
1	3	455	Installation Problem	7
2	4	940	Installation Problem	5
3	5	1744	Installation Problem	8
4	4	621	Installation Problem	7
5	6	621	Installation Problem	7
6	2	1831	Installation Problem	7
7	4	1198	Installation Problem	6
8	3	1703	Installation Problem	6
9	5	1703	Installation Problem	6
10	2	1449	Installation Problem	6
11	2	1288	Installation Problem	5
12	6	493	Installation Problem	5
13	6	766	Installation Problem	7
14	5	766	Installation Problem	7
15	6	1754	Installation Problem	7

Rows: 86 Expand

Unknown integration DataFrame as s

```
SELECT
    column_name,
    data_type,
    STRING_AGG(CONCAT(table_schema, '.', table_name), ', ') AS table_list
FROM information_schema.columns
WHERE table_schema = 'public'
  AND table_name IN ('survey', 'support')
GROUP BY column_name, data_type
HAVING COUNT(*) > 1;
```

	col...	d	table_list
0	customer_id	integer	public.survey, public.support
1	id	integer	public.support, public.survey

Rows: 2 Expand