Python Summer Party Challenge

by Interview Master

Day 1 of 15

WhatsApp

You are a Product Analyst on the WhatsApp team investigating group messaging dynamics. Your team wants to understand how large groups are being used and their messaging patterns. You'll leverage data to uncover insights about group participation and communication behaviors.

Challenge Questions

Q1:

What is the maximum number of participants among WhatsApp groups that were created in October 2024? This metric will help us understand the largest group size available.

Q2:

What is the average number of participants in WhatsApp groups that were created in October 2024? This number will indicate the typical group size and inform our group messaging feature considerations.

Q3:

For WhatsApp groups with more than 50 participants that were created in October 2024, what is the average number of messages sent? This insight will help assess engagement in larger groups and support recommendations for group messaging features.

My Solution - Q1

Day 1 Python Challenge

```python
# Note: pandas and numpy are already imported as pd and np
# The following tables are loaded as pandas DataFrames with th
e same names: dim_groups
# Please print your final result or dataframe


october_groups = dim_groups[
    (dim_groups['created_date'] >= "2024-10-01") &
    (dim_groups['created_date'] < "2024-11-01")
]
max_participants = october_groups['participant_count'].max()
print("Maximum participants in October 2024 groups:", max_part
icipants)
```

InterviewMaster.AI

My Solution - Q2

Day 1 Python Challenge

```python
import pandas as pd
dim_groups['created_date'] = pd.to_datetime(dim_groups['created_date'])
october_df = dim_groups[
    (dim_groups['created_date'] >= "2024-10-01") &
    (dim_groups['created_date'] < "2024-11-01")
]

avg_participants = october_df['participant_count'].mean()

print("Average participants in October 2024 groups:", avg_participants)
```

My Solution - Q3

Day 1 Python Challenge

```python
import pandas as pd
dim_groups['created_date'] = pd.to_datetime(dim_groups['created_date'])

october_group = dim_groups[
  (dim_groups['created_date'] >= "2024-10-01") &
  (dim_groups['created_date'] < "2024-11-01") &
  (dim_groups['participant_count'] > 50)
]

# Calculate average number of messages
avg_messages = october_group['total_messages'].mean()

print("Average messages in large October 2024 groups:", avg_messages)
```