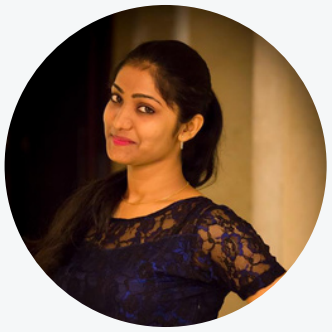

Avantika Penumarty

📢 Day 2/30 - Data Engineering Challenge 🚀

Difficulty Level: Easy

Today, we're building on yesterday’s fundamentals and pushing a bit further! Let’s test your SQL , Python , ETL transformations, and data modeling concepts

SQL Challenge - Day 2/30

@zero2dataengineer

🚀 Concept Focus: Window Functions & Ranking

📌 Topic: Using ROW_NUMBER() to rank employee salaries

❓ What will be the output of this SQL query?

Employee Salary Ranking Query

```sql
SELECT name, salary,
       ROW_NUMBER() OVER (PARTITION BY department ORDER BY salary DESC) as rank
FROM employees;
```

- A) Assigns a rank to each employee within a department
- B) Counts total employees
- C) Generates unique numbers across all employees
- D) SQL Error

💡 Hint: Window functions allow row ranking within partitions, useful for leaderboard-style queries.

🧐 Drop your answer below!

📢 Solutions available on Substack 👉 Subscribe for free

https://zero2dataengineer.substack.com/subscribe

Python Challenge - Day 2/30

@zero2dataengineer

🚀 Concept Focus: Dictionary Operations & Mutability
📌 Topic: Understanding how Python dictionaries behave when assigned to another variable

❓ What will be the output of this Python code?

Dictionary Modification

```
dict_1 = {'a': 1, 'b': 2}
dict_2 = dict_1
dict_2['c'] = 3
print(dict_1)
```

❓ What will be the output?

- A) {'a': 1, 'b': 2}
- B) {'a': 1, 'b': 2, 'c': 3}
- C) {'c': 3}
- D) Error

💡 Hint: Dictionaries, like lists, are mutable and passed by reference. Changes in dict_2 affect dict_1.

🧐 Drop your answer below!
📢 Solutions available on Substack 👉 Subscribe for free

https://zero2dataengineer.substack.com/subscribe

⚡ ETL Challenge - Day 2/30

@zero2dataengineer

📌 Concept Focus: Data Cleaning & Transformation
🔎 Topic: Identifying and handling duplicate records in an ETL pipeline

❓ What is the best way to remove duplicates from a dataset before loading it into a data warehouse?

🔘 A) Use DISTINCT in SQL queries
🔘 B) Apply a GROUP BY operation with aggregation
🔘 C) Deduplicate at the source system before ingestion
🔘 D) All of the above

💡 Hint: Different use cases need different deduplication strategies!

🧐 Drop your answer below!
📢 Solutions available on Substack 👉 Subscribe for free

https://zero2dataengineer.substack.com/subscribe

Data Modeling - Day 2/30

@zero2dataengineer

📌 Concept Focus: Many-to-Many Relationships
🔎 Topic: How do you properly model a many-to-many relationship in a relational database?

❓ Which approach correctly models a many-to-many relationship between students and courses?

- 🔘 A) Adding a course_id column in the students table
- 🔘 B) Adding a student_id column in the courses table
- 🔘 C) Creating a student_courses junction table
- 🔘 D) Using only foreign keys without a separate table

💡 Hint: Think about how one student can enroll in multiple courses & vice versa!

🧐 Drop your answer below!
📢 Solutions available on Substack 👉 Subscribe for free

https://zero2dataengineer.substack.com/subscribe

🚀 Level Up Your Data Engineering Skills!

Enjoying these challenges?

Get 30 days of FREE challenges before we move to exclusive paid deep dives!

- **Upgrade today to unlock monthly/Annual Plan for more Deep Dives : https://zero2dataengineer.substack.com/subscribe**

✅ Advanced breakdowns
✅ Live runnable SQL & Python code
✅ Real-world DE interview strategies
📌 Upgrade now for Full Access! 🚀