

[Google Discussions](#)

### Exam Professional Data Engineer All Questions

View all questions & answers for the Professional Data Engineer exam

[Go to Exam](#)

## EXAM PROFESSIONAL DATA ENGINEER TOPIC 1 QUESTION 180 DISCUSSION

Actual exam question from Google's Professional Data Engineer

Question #: 180

Topic #: 1

[\[All Professional Data Engineer Questions\]](#)

You are migrating an application that tracks library books and information about each book, such as author or year published, from an on-premises data warehouse to BigQuery. In your current relational database, the author information is kept in a separate table and joined to the book information on a common key. Based on Google's recommended practice for schema design, how would you structure the data to ensure optimal speed of queries about the author of each book that has been borrowed?

- A. Keep the schema the same, maintain the different tables for the book and each of the attributes, and query as you are doing today.
- B. Create a table that is wide and includes a column for each attribute, including the author's first name, last name, date of birth, etc.
- C. Create a table that includes information about the books and authors, but nest the author fields inside the author column.
- D. Keep the schema the same, create a view that joins all of the tables, and always query the view.

[Show Suggested Answer](#)

by [AWSandeep](#) at *Sept. 2, 2022, 9:22 p.m.*

### Comments

[Submit](#)

🗄️ 👤 **musumusu** Highly Voted 8 months, 2 weeks ago

C

if data is time based or sequential, find partition and cluster option

if data is not time based,

always look for denormalize / nesting option.

👍 ↩️ 🚩 upvoted 10 times

🗄️ 👤 **zelck** Highly Voted 11 months, 1 week ago

**Selected Answer: C**

C is the answer.

<https://cloud.google.com/bigquery/docs/best-practices-performance-nested>

Best practice: Use nested and repeated fields to denormalize data storage and increase query performance.

Denormalization is a common strategy for increasing read performance for relational datasets that were previously normalized. The recommended way to denormalize data in BigQuery is to use nested and repeated fields. It's best to use this strategy when the relationships are hierarchical and frequently queried together, such as in parent-child relationships.

👍 ↩️ 🚩 upvoted 5 times

🗄️ 👤 **AzureDP900** Most Recent 10 months, 1 week ago

C. Create a table that includes information about the books and authors, but nest the author fields inside the author column.

👍 ↩️ 🚩 upvoted 1 times

🗄️ 👤 **Atnafu** 11 months, 2 weeks ago

C

Best practice: Use nested and repeated fields to denormalize data storage and increase query performance.

👍 ↩️ 🚩 upvoted 2 times

🗄️ 👤 **dish11dish** 11 months, 3 weeks ago

**Selected Answer: C**

Use nested and repeated fields to denormalize data storage which will increase query performance. BigQuery doesn't require a completely flat denormalization. You can use nested and repeated fields to maintain relationships

👍 ↩️ 🚩 upvoted 2 times

🗄️ 👤 **Thobm** 1 year, 1 month ago

**Selected Answer: C**

<https://cloud.google.com/bigquery/docs/best-practices-performance-nested>

👍 ↩️ 🚩 upvoted 1 times

🗄️ 👤 **ducc** 1 year, 2 months ago

**Selected Answer: C**

C is correct

👍 ↩️ 🚩 upvoted 2 times

🗄️ 👤 **AWSandeep** 1 year, 2 months ago

**Selected Answer: C**

C. Create a table that includes information about the books and authors, but nest the author fields inside the author column.

👍 ↩️ 🚩 upvoted 2 times

# EXAMTOPICS

Platform

> Home

> Examtopics PRO

> All Exams

> Training Courses

