

What are **two** benefits of storing this data in MongoDB with JSON over a relational database management system such as Postgres? Please reference specific examples from the **business** collection to back up your claims.

Format your answer as follows:

1. Benefit #1, Example #1.
2. Benefit #2, Example #2.

**Limit each benefit to one sentence and each example to one sentence for a total of at most four sentences.**

Benefit #1: MongoDB allows flexible schema design, so documents can include fields only when relevant. Example #1: The attributes field in “Oskar Blues Taproom” contains detailed subfields like BusinessParking, HappyHour, and WiFi, which wouldn’t fit neatly into fixed relational tables. Benefit #2: MongoDB supports rich embedded documents, which enables easy storage of complex nested data without normalization. Example #2: The hours field in both businesses stores each day’s opening time in a nested dictionary, avoiding the need for a separate “hours” table as would be needed in Postgres.



---

### 0.0.1 Question 2d

In the last question, you performed equivalent left joins in both Postgres and Mongo. Now, examine their query plans, paying special attention to `executionTimeMillis`. Which database system was faster? What gives the database system you chose an advantage over the other? Keep your response to at most three sentences.

Postgres was faster, with an execution time of 3.924 ms compared to Mongo's 400 ms. This is likely because Postgres uses a hash join with indexed scan, which is optimized for relational joins. MongoDB's \$lookup performs a left outer join but requires scanning and matching across documents, which can be slower without a proper index.



1. What do you notice about how the columns of `business_df` are constructed, e.g. how are fields and subfields represented in the dataframe?
2. How are values that are not found in every document handled in the pandas dataframe?

Hint: You will need to horizontally scroll to view all the column names and values.

**Format your answer as follows and use one sentence per question:**

1. Sentence 1
2. Sentence 2

```
In [ ]: 1)The columns of business_df are flattened so that nested fields and subfields are split into s
        2)Values that are not found in every document are represented as NaN in the Pandas dataframe, in
```

