**1. NoSQL Model Choice: MongoDB**

**Justification:**

* **Query Requirements:** All 7 exam queries require complex aggregations and filtering on multiple fields
* **Data Structure:** Semi-structured data with nested objects (host info, location, reviews) fits document model perfectly
* **Scalability:** Horizontal scaling available if dataset grows

**Why NOT Redis or HBase:**

* Redis: Would require complex secondary structures for aggregations
* Hbase: We can use it, but I prefer MongoDb instead

**2. Schema Design**

**Collection:** listings (single collection approach)

json

{

"\_id": ObjectId,

"listing\_id": int,

"name": string,

"property\_type": string

"host": {

"host\_id": int,

"host\_name": string

"is\_superhost": boolean

},

"location": {

"city": string,

"neighbourhood": string,

"latitude": float,

"longitude": float

},

"reviews": {

"count": int

},

"instant\_bookable": boolean,

"last\_scraped": datetime

}

**Design Decisions:**

* **Denormalization:** Host info embedded in listing (no joins needed)
* **Nested Objects:** Logical grouping of related fields
* **Indexes:** Created on all query-critical fields

**3. Data Quality & Cleaning**

**Cleaning Operations:**

* **Boolean conversion: 't'/'f' → true/false**
* **Date parsing: String dates → Date**
* **Column selection: Kept only relevant fields**

**4. Migration Results**

**A screenshot of a computer

AI-generated content may be incorrect.**

**5. Query Results**

**=== QUERY 1: Listings per property type ===**

**Entire rental unit: 80516 listings**

**Private room in rental unit: 5980 listings**

**Entire condo: 2679 listings**

**Room in boutique hotel: 1256 listings**

**Entire loft: 1007 listings**

**Room in hotel: 1003 listings**

**Entire home: 696 listings**

**Private room in bed and breakfast: 590 listings**

**Private room in condo: 344 listings**

**Entire serviced apartment: 277 listings**

**Shared room in rental unit: 274 listings**

**Entire townhouse: 268 listings**

**Private room in home: 122 listings**

**Private room in townhouse: 86 listings**

**Private room in loft: 73 listings**

**Room in aparthotel: 73 listings**

**Shared room in hostel: 54 listings**

**Private room in hostel: 53 listings**

**Entire place: 46 listings**

**Private room in guesthouse: 45 listings**

**Private room in guest suite: 38 listings**

**Private room: 37 listings**

**Entire vacation home: 37 listings**

**Private room in casa particular: 36 listings**

**Entire guesthouse: 28 listings**

**Tiny home: 26 listings**

**Shared room in condo: 23 listings**

**Entire guest suite: 22 listings**

**Room in serviced apartment: 19 listings**

**Entire villa: 19 listings**

**Shared room in hotel: 17 listings**

**Private room in serviced apartment: 17 listings**

**Entire home/apt: 16 listings**

**Room in bed and breakfast: 11 listings**

**Boat: 10 listings**

**Private room in boat: 9 listings**

**Shared room in bed and breakfast: 6 listings**

**Room in hostel: 6 listings**

**Shared room in loft: 6 listings**

**Houseboat: 6 listings**

**Private room in tiny home: 5 listings**

**Casa particular: 4 listings**

**Shared room in guesthouse: 4 listings**

**Camper/RV: 4 listings**

**Shared room in townhouse: 3 listings**

**Private room in vacation home: 3 listings**

**Shared room in boutique hotel: 3 listings**

**Shared room in home: 3 listings**

**Room in rental unit: 2 listings**

**Cave: 2 listings**

**Shared room in tiny home: 2 listings**

**Dome: 1 listings**

**Entire bed and breakfast: 1 listings**

**Private room in villa: 1 listings**

**Shared room in serviced apartment: 1 listings**

**Shared room: 1 listings**

**Castle: 1 listings**

**Shared room in ice dome: 1 listings**

**Shared room in guest suite: 1 listings**

**Entire bungalow: 1 listings**

**Earthen home: 1 listings**

**Shared room in cabin: 1 listings**

**Tower: 1 listings**

**Private room in cave: 1 listings**

**Barn: 1 listings**

**Floor: 1 listings**

**=== QUERY 2: Listings on June 12, 2024 ===**

**Listings on June 12, 2024: 41323**

**=== QUERY 3: Top 5 listings by reviews ===**

**1. Listing 17222007: 3067 reviews**

**2. Listing 26244787: 2620 reviews**

**3. Listing 41020735: 2294 reviews**

**4. Listing 40194697: 2105 reviews**

**5. Listing 35145338: 2048 reviews**

**=== QUERY 4: Total unique hosts ===**

**Total unique hosts: 71979**

**=== QUERY 5: Instant bookable rentals ===**

**Instant bookable: 22094**

**Total listings: 95885**

**Percentage: 23.0%**

**=== QUERY 6: Hosts with 100+ listings ===**

**Hosts with 100+ listings: 24**

**Percentage of total hosts: 0.03%**

**Host 314994947 (Blueground): 730 listings**

**Host 33889201 (Veeve): 497 listings**

**Host 50502817 (Pierre De WeHost): 426 listings**

**Host 50978178 (Sébastien): 307 listings**

**Host 26981054 (Cédric De ClickYourFlat): 274 listings**

**Host 460047164 (FlexLiving): 232 listings**

**Host 7642792 (Ludovic): 211 listings**

**Host 436103373 (David Et Warren): 210 listings**

**Host 528015349 (Checkmyguest): 189 listings**

**Host 125797498 (Welkeys): 170 listings**

**Host 335998296 (Studioprestige): 154 listings**

**Host 51567288 (Sweet Inn): 149 listings**

**Host 1112584 (IntoParis): 145 listings**

**Host 564251645 (Rusard): 137 listings**

**Host 440295601 (Parisian Home): 132 listings**

**Host 517515174 (Barnes): 122 listings**

**Host 99040006 (Jérémy): 120 listings**

**Host 28313443 (Michael & Johanna): 119 listings**

**Host 24495283 (Giacomo): 111 listings**

**Host 506389460 (Check My Guest): 109 listings**

**Host 374552379 (Checkmyguest): 104 listings**

**Host 512746089 (Checkmyguest): 104 listings**

**Host 499962530 (Check My Guest): 103 listings**

**Host 21630783 (Pierre): 102 listings**

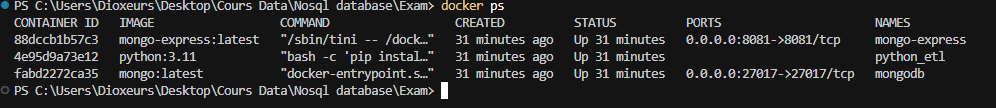
**=== QUERY 7: Superhosts ===**

**Unique superhosts: 10088**

**Total hosts: 71979**

**Percentage: 14.0%**

**6. Docker Setup**

****