**Amazon Data Engineer Example Questions**

Let’s get into the primary categories of questions you’ll encounter in the Amazon data engineer interview.

Here are the most common question categories based on the questions reported by Amazon

DE candidates on Glassdoor.

 Behavioral interview questions

 SQL interview questions

 Data management interview questions

 Data modeling

 Data warehousing

 Data pipelines

 Coding interview questions

 Others: analytical problem-solving, data visualization

 Do note that each Amazon business unit or team will have different technical requirements. Your recruiter should be able to advise you on which aspects to focus on.

**1. SQL Interview Questions**

**Why It Matters:**

Amazon’s backend processes billions of data operations daily. Data engineers must demonstrate proficiency in handling vast volumes of structured and unstructured data. SQL assessments are integral to evaluating this ability.

**What to Expect:**

You'll solve problems that mimic real-world scenarios, including:

 Writing efficient SQL queries using window functions, subqueries, joins, and aggregate functions.

 Demonstrating performance tuning and query optimization techniques.

 Applying SQL skills to data modeling and warehousing problems.

**Example SQL Questions:**

1. **Peak Activity:**

 *Question*: Given a dataset, find the time period when the most people were online, measured in seconds.

 *Focus*: Time-series analysis, window functions, and efficient query writing.

2. **Employee Data Mining:**

 *Question*: From a table with columns datetime, employee, and customer\_response (free text), extract phone numbers embedded in customer\_response and find the top 10 employees associated with the most numbers.

 *Focus*: String manipulation and ranking techniques.

3. **Sales Trends Analysis:**

 *Question*: Display all products with more than 50% sales increase from the previous month to the current month.

 *Focus*: Analytical queries involving percentage calculations and time-based aggregation.

4. **Retention Analysis:**

 *Question*: Calculate the first-day retention rate for players. Retention is defined as players logging in on the second day after their initial login.

 *Focus*: Date calculations and user behavior analysis.

5. **Subscriber Campaign Data:**

 *Question*: Write a query to identify inactive subscribers for a marketing re- engagement campaign.

 *Focus*: Filtering and conditional logic.

6. **Revenue Calculation:**

 *Question*: Write a query to get the total revenue generated by each subscriber in 2014.

 *Focus*: Aggregation and filtering by date ranges.

7. **Performance Troubleshooting:**

 *Scenario*: Given a specific SQL error or performance degradation, explain the cause and propose fixes.

**Tips for SQL Success:**

 Practice using complex joins, CTEs, and window functions.

 Familiarize yourself with query optimization techniques (e.g., indexing, analyzing execution plans).

 Be comfortable explaining your approach during the interview.

**2. Data Management Questions**

**Why It Matters:**

With its vast datasets, Amazon relies on efficient data modeling, storage, and retrieval solutions. You'll need to design schemas, build pipelines, and solve data warehousing challenges.

**Data Modeling Questions**

1. How do you create a schema that tracks customer addresses, including address changes over time?

 Focus: Temporal data management and normalization.

2. Design a data model to track a product’s journey from vendor to warehouse to delivery.

 Focus: Workflow-based schema design.

3. Should we apply normalization rules to a star schema? Why or why not?

 Focus: Trade-offs in schema design for analytical workloads.

4. What’s a chasm trap? How do you avoid it?

 Focus: Advanced relational database concepts.

**Data Warehousing Questions**

1. Design a data warehouse to capture sales and assist the customer support team with ticket management.

 Focus: Dimensional modeling and schema design.

2. Provide a schema for an OLAP system tailored for Amazon's retail data.

 Focus: Analytical query optimization.

**Data Pipeline Questions**

1. Given a dataset and increasing data volume, how would you design an ETL pipeline using AWS products?

 Focus: Scalable pipeline architecture, fault tolerance, and cloud services.

2. Write a script to transform raw logs into a structured format for analytics.

 Focus: End-to-end ETL design and implementation.

**3. Coding Interview Questions**

**Why It Matters:**

Amazon’s data engineers solve key challenges with efficient, scalable code. These questions test your logical thinking, data manipulation skills, and ability to work with data structures.

**What to Expect:**

Most coding questions are in Python, focusing on:

 Data structure manipulation (e.g., dictionaries, lists).

 Algorithmic problem-solving.

 File and data processing.

**Example Coding Questions:**

1. **Odd Number Sorting:**

 *Question*: Write a function to sort an array, returning only odd numbers.

2. **Unique Values Preservation:**

 *Question*: Find non-duplicate numbers from a list while preserving the original order. Example: [1,1,3,2,5,6,5] → [1,3,2,5,6].

3. **Maximum Occurrences:**

 *Question*: Given a list, return the numbers with the highest count.

4. **JSON Flattening:**

 *Question*: Write a function to flatten nested JSON objects into a single key- value dictionary.

5. **Array Pair Sum:**

 *Question*: Write code to find two numbers in an array that sum up to x.

6. **Stack Implementation:**

 *Question*: Implement a stack using a linked list.

**4. Analytical Problem-Solving and Data Visualization**

**Why It Matters:**

Amazon values data-driven insights and clear communication of findings. These questions test your ability to analyze data, create metrics, and visualize trends.

**Key Questions:**

1. How have you used statistics in data analysis? Provide examples.

2. Describe a dashboard or visualization project you’ve completed. What tools did you use, and how did you define KPIs?

3. What’s your approach to building end-to-end dashboards?

**5. Behavioural interview questions**

**Customer Obsession interview questions**

 **Definition**: Leaders prioritize customers' needs and work backward to enhance their experience. They earn trust and relentlessly focus on delivering value.

 **Why it matters**: Amazon emphasizes customer-centric innovation over competitor analysis, making this the cornerstone of its culture.

 **How to prepare**:

1. Highlight instances where you prioritized customer needs, even in challenging scenarios.

2. Share specific examples demonstrating empathy, problem-solving, and customer trust-building.

 **Example Questions**:

 Tell me about a time you went above and beyond for a customer.

 How did you handle a challenging customer interaction?

 Can you share a situation where you simplified a process for customers?