**WALMART DATA ENGINEER PREP TIPS**

Linked list palindrome

How Spark works and its internals

ETL tools that you used

technical business focused questions

given: [126,125,67,28, 658] expected: [9, 8, 13, 10, 19] without using a 2 loops

What is a case class in Scala

Merging two arrays without using a third array

SQL queries

some string manipulation

spark, Scala and SQL, and some system design basics.

SQL query optimization. Relational vs non-relational databases

Questions was about my previous work, data warehousing, SQL and shell scripting.

how to promote unsold products

//modified LFU cache to get the unsold products then used many to many between products and users

//to get the users and notify them regarding the discounts/offers on unsold products.

//caching algorithm in which the Least Frequently Used item in the cache is removed whenever the cache's

//capacity limit is reached. This means that for every item in our cache we have to keep track of how frequently it is used.

how to see if there’s a gender gap based on a couple tables

Spark Scala hive MapReduce java database exceptional questions

Enhance your theoretical knowledge.

Coding skills using collections

Personality Questions:

1. Why do you want to work for Walmart?

2. If you were told that you need to work as quickly as possible, how do you make sure that the quality of your work doesn’t suffer?

3. Can you describe a time you went above and beyond in your projects?

4. Do you prefer working as part of a team or independently?

5. On a scale of one to ten, how would you rate your communication skills? Why would you give it that score?

6. How do you design a Supermarket for older people?

7. If you are opening a new Walmart branch, how would you decide how many cash registers/(or any) are needed for the store?

8. How would you keep Developers working on a product motivated and turning out quality work?

9. Why do you want to be a Data Engineer?

10. Where do you see yourself in 5 years?

Walmart Specific:

https://www.youtube.com/watch?v=WARzZVfT4lU

Walmart Labs Data Engineer Interview!

Check the description too

behavioral \*Walmart Specific\*

https://www.youtube.com/watch?v=TCDhY69Xjgc

Top 20 Walmart Interview Questions and Answers for 2022

https://www.youtube.com/watch?v=k5NhuxCwrok

WALMART Interview Questions & Answers 2020! (Walmart Interview Process, Tips and ANSWERS!)

https://articles.bplans.com/hiring-employees-here-are-over-70-must-ask-interview-questions/

https://blog.insightdatascience.com/the-data-engineering-ecosystem-in-2017-2c2a3429350e

<https://github.com/andkret/Cookbook/blob/master/sections/01-Introduction.md#data-engineer>

1). 49 minutes

2). I was not. One of the interviewers asked me a battery of question about Kafka and I was not able to answer all of them

3). Projects, exclusively Spark and Kafka.

Questions:

1. How would you keep stock market data that an end user could would interact with. Answer: Bi-Temporal Database //A temporal database stores data relating to time instances. It offers temporal data types and stores information relating to past, present and future time.Bitemporal database design is a method of storing time dependant data records to represent both the history of the facts and the history of changes to the records in the database. Bitemporal databases permit queries over two orthogonal time dimensions: valid time and transaction time.
2. If you noticed that your Kafka consumer was only retrieving data from one partition despite there being 3, how would you trouble shoot this? Answer: I would first check to make sure the brokers where the other partitions are stored are up, running, and connected to the cluster. If they were, then I would check to see how the message keys are being created since messages are designated to partitions based upon keys.
3. What does the 'auto.offset.reset' configuration used for in Kafka. Answer: to define the behavior of the consumer when there is no committed position (which would be the case when the group is first initialized) or when an offset is out of range. You can choose either to reset the position to the “earliest” offset or the “latest” offset (the default).to define the behavior of the consumer when there is no committed position (which would be the case when the group is first initialized) or when an offset is out of range. You can choose either to reset the position to the “earliest” offset or the “latest” offset (the default).
4. If you notice that your offset is skipping what would cause that? How would you debug that? Answer: I would first check to ensure that whatever medium I'm using to monitor the offset is working properly because there's a chance that the consumer isn't skipping, it just seems like it is

// https://stackoverflow.com/questions/70942159/kafka-consumer-how-to-recognized-offset-skipping-missing-offsets