Walmart Data Science Interview

Walmart can process roughly 2.5 petabytes of data every hour.

Vimarsh Karbhari   Follow

Jun 19, 2018 · 3 min read

To make sense of all of this information the company has created what it calls its Data Café — a state-of-the-art analytics hub located within its Bentonville, Arkansas headquarters. To say that the sheer scale of this operation is gigantic is an understatement.
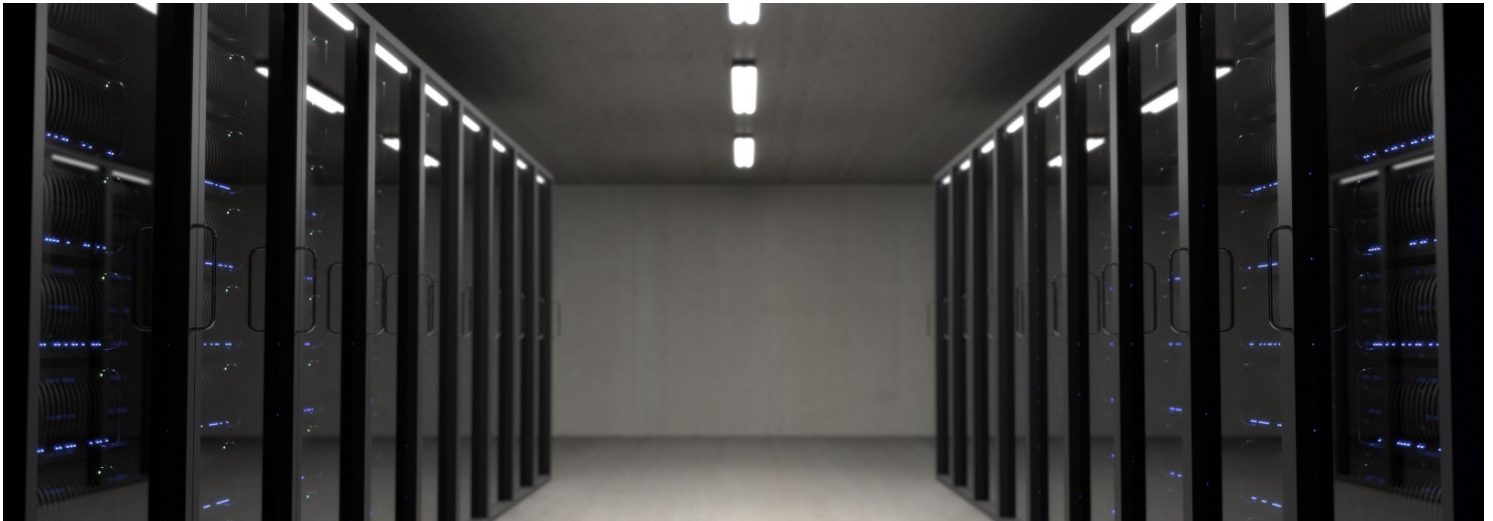

Photo by Manuel Geissinger from Pexels

There are 200 billion rows of transactional data (representing the past few weeks), the Data Café pulls in information from 200 sources, including meteorological data, economic data, Nielsen data, telecom data, social media data, gas prices, and local events databases. Anything within these vast and varied datasets could hold the key to the solution to a particular problem, and Walmart's algorithms are designed to run through them in microseconds and present the data in a consumable way.

Interview Process

The interview process usually consists of one phone screen, one onsite interview (4 technical + 1 lunch interview). One of the technical interviews is coding, one is on probability and statistics and big data questions, one is on machine learning. Usually there is also a technical interview which is with hiring manager which involves some basic coding.

Walmart Labs Office Locations Source: https://www.walmartlabs.com/

Important Reading

1. Better Data for eRetail: Walmart Labs Presentation

2. All Data Science related blog articles: Walmart Labs Data Science Articles

3. How Walmart uses Kafka with Data Science: Kafka Ecosystem on Walmart's clouds

AI/Data Science Related Questions

- What is the difference between Gradient Boosting and Random Forest?

- What are the different types of Sampling methods that you have used?

- Explain the difference between bagged and boosting models.

- What is cross entropy?

- What is multi-collinearity, how do you fix it in a regression?

- What is the significance of log odds?

- What is the relationship between sample size and margin of error?

- What are type I and II errors?

- What is F distribution?

- Define parametric and non-parametric methods. Give some examples.

- What are the data structures in Python? What are the data structures in the Pandas package in Python?

- What are generator, iterators, list comprehension in Python?

- Give some problems or scenarios where map-reduce concept works well and where it doesn't work.

- How can you design a product recommendation system based on taxonomy?

- Given a dataset having employee id and manager id find the employees who are also managers ?

- A person is using search engine to find something, you know nothing about her/him, how do you come up with am algorithm that will predict what she/he needs after the user types only a few letters ?

- Print all the combinations in a binary tree. Use any language of your choice.

- Given a sheet of data provide explanations on how you would make analyses such as linear regression and decision trees.

- How would you deal with an older professional who consulted you but does not believe in your analytical results and sticks to his older methods?

- Write an algorithm to know if a String is palindrome using Linked list.

Reflecting on the Questions

Walmart has huge amounts of data in the E-commerce/retail world. Their Data Science team has been around from time before Data Science became so hyped. Questions reflect the same sense of maturity. They use MapReduce, Hive, HDFS and Spark internally. Questions and interviews consist of a mixture of coding and model based questions which align towards well rounded data science and technical skills. A little bit of practice can surely land you in a job with the world's largest retailer.

Consumable list: Walmart Data Science Interview Questions

My previous article was featured on KDnuggets:
https://www.kdnuggets.com/2018/06/netflix-data-science-interview-questions-acing-the-ai-interview.html

Thanks for reading! 😊 *If you enjoyed it, test how many times can you hit* 👏 *in 5 seconds. It's great cardio for your fingers AND will help other people see the story.*

· · ·

Subscribe to our newsletter here. We are building a new course to help people ace data science interviews. Sign up below to join the wait-list!

Newsletter

Subscribe to the Acing AI/Data Science Newsletter. It is FREE! Reducing the entropy in data science. Helping you with…

www.acingdatascienceinterviews.com

Interview   Data Science   Artificial Intelligence   Machine Learning   Deep Learning

About   Help   Legal

Get the Medium app