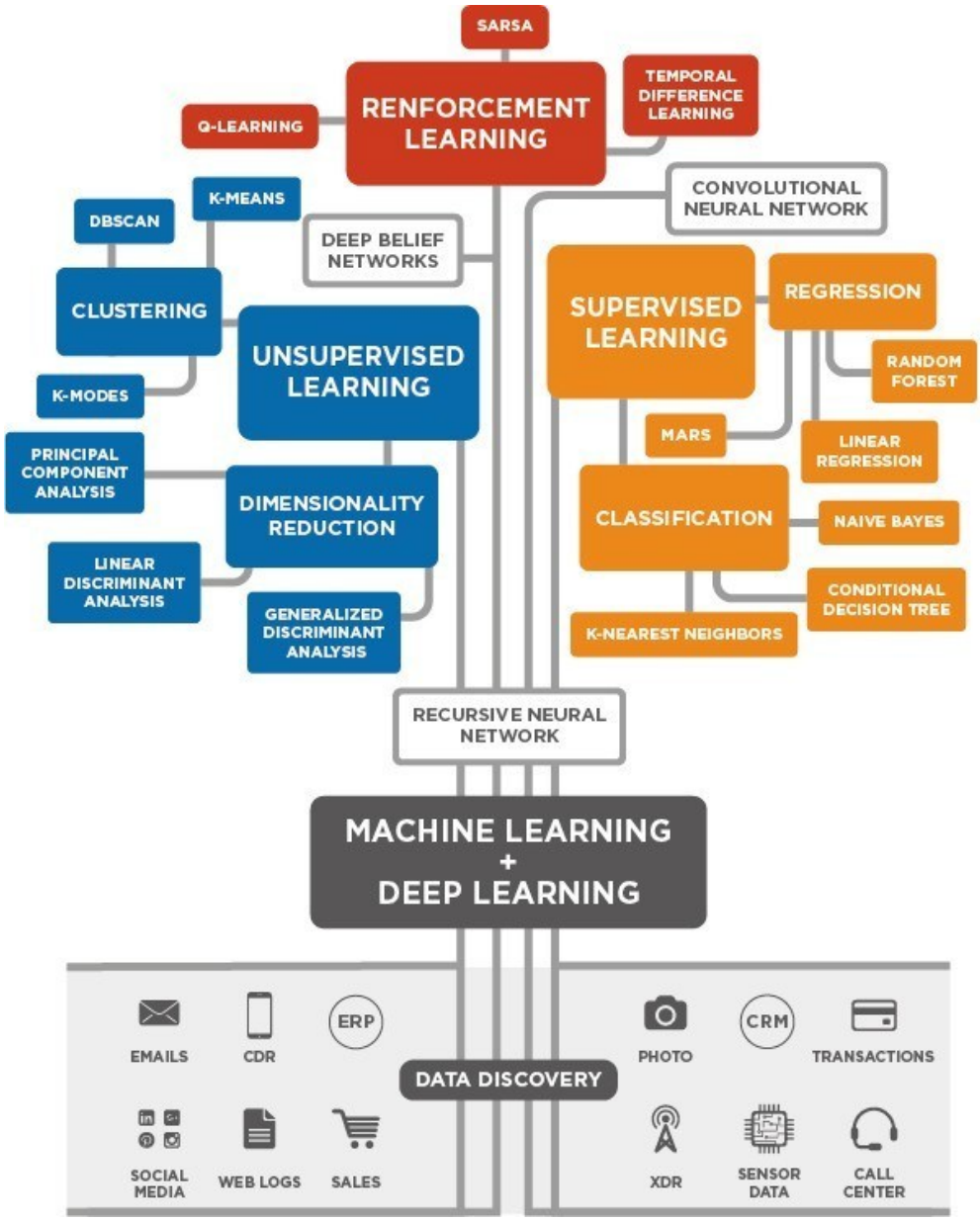
THE DATA SCIENCE INTERVIEW STUDY GUIDE

By Ben Rogojan Data science technical interviews, like other technical interviews require plenty of preparation. There are lots of subjects that need to be covered in order to ensure you...



by Ben R. *May 20, 2019*

By Ben Rogojan

Data science technical interviews, like other technical interviews require plenty of preparation. There are lots of subjects that need to be covered in order to ensure you are ready for a data science interview.

Before we get started. We want to point out some tips.

One of the things we have noticed is there are several types of data science interviews.

Some data science interviews are very product and metric driven. These interviews focus more on asking product questions like what kind of metrics would you use to show what you should improve in a product. These are often paired with SQL and some Python questions.

The other type of data science interview tends to be a mix of programming and machine learning.

We recommend asking the recruiter if they don't provide you any tips. Some companies are very good at keeping interviews consistent but even then, teams sometimes deviate depending on what they are looking for. Here are some examples of what we have noticed about some companies data science interviews.

Airbnb — Product Heavy, Metrics diagnostics, Metrics creation, A/B testing, Tons of behavioral questions and take home.

Netflix — Product-sense questions, A/B testing, experimental design ,metric design

Microsoft — Programming Heavy, Binary Tree Traversal, SQL ,Machine Learning

Expedia — Product, Programming, sql, product sense, Machine learning questions about SVM, regression and decision tree

In order to keep track of your progress we have put together a study checklist to help you out!

Download The Data Science Interview Checklist

Let's first start with making sure you can explain the basic data science algorithms.

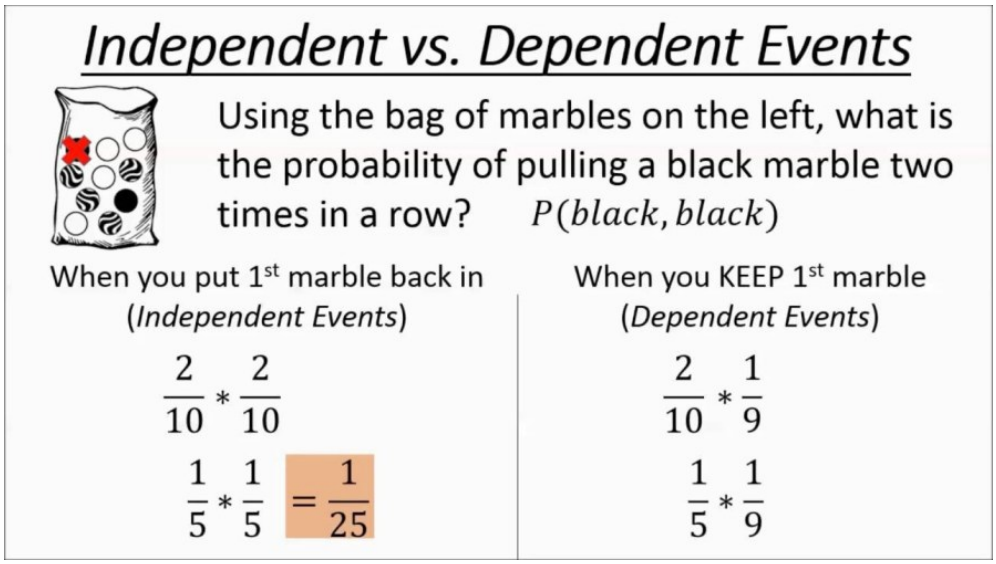
Machine Learning Algorithms

1. Logistic Regression—Video
2. A/B Testing?—Video
3. Decision Tree—Post
4. SVM—Post
5. How SVM—Video
6. Principal Component Analysis: PCA—post
7. Principal Component Analysis—Video

8. Adaboost—Post
9. AdaBoost—Video
10. A Gentle Introduction to the Gradient Boosting Algorithm for Machine Learning—Post
11. Gradient Boost Part 1: Regression Main Ideas—Video
12. K-Means Clustering—The Math of Intelligence—Video
13. Bayesian Network—Post
14. Neural Network—Post
15. Dimensionality reduction algorithms—Post
16. How kNN algorithm works—Video

Probability And Statistics

A common question you might get at FAANG companies and other tech companies alike is the occasional probability or statistics question. The questions won't necessarily require complex math. However, if you haven't thought about independent and dependent probabilities in while. It is good to review setting up the basic formulas.



Probability Videos

1. Dependent probability introduction
2. Independent & dependent probability
3. Independent Problems
4. Conditional Prob Article

Probability Quiz

1. Probability & Statistics—Set 6
2. Probability & Statistics—Set 2
3. Independent Probability
4. Dependent Probability

Probability Interview Questions

Most of these questions are either similar to the ones we have been asked or taken directly from glassdoor.com.

1. A die is rolled twice. What is the probability of showing a 3 on the first roll and an odd number on the second roll?
2. In any 15-minute interval, there is a 20% probability that you will see at least one shooting star. What is the probability that you see at least one shooting star in the period of an hour?
3. Alice has 2 kids and one of them is a girl. What is the probability that the other child is also a girl? You can assume that there are an equal number of males and females in the world.
4. You're about to get on a plane to Seattle. You want to know
5. How many ways can you split 12 people into 3 teams of 4?

Statistics Pre-Quizzes

1. Data Science Probability Statistics 14

Statistics Concepts

Statistics is a broad concept so don't get too bogged down in the details of each of these videos. Instead, just make sure you can explain each of these concepts at the surface level.

1. Bias Variance Trade Off
2. Confusion Matrix
3. ROC curve
4. Normal Distribution
5. P-Value
6. Pearson Spearman

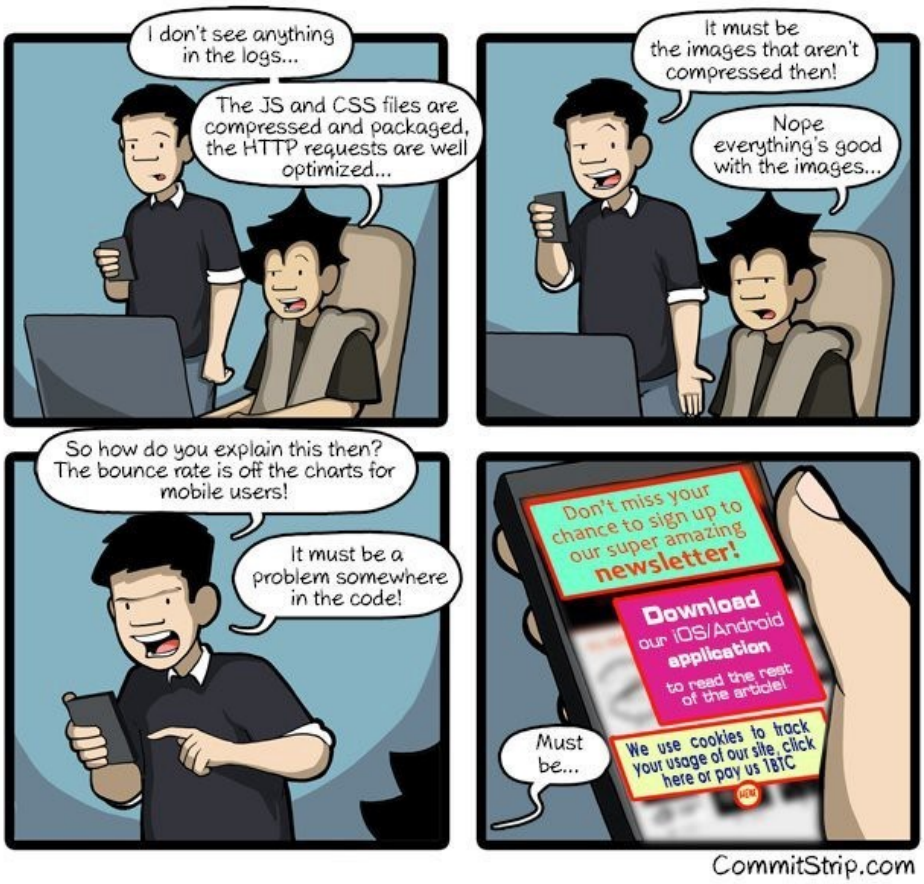
1. Normal distribution problem: z-scores (from ck12.org)
2. Continuous Probability Distributions
3. Standardizing Normally Distributed Random Variables (fast version)
4. Statistics 101: Simple Linear Regression, The Very Basics
5. Statistics 101: Linear Regression, Outliers and Influential Observations
6. Statistics 101: ANOVA, A Visual Introduction
7. Statistics 101: Multiple Regression, The Very Basics
8. Statistics: Variance of a population | Probability and Statistics | Khan Academy
9. Expected Value: E(X)
10. Law of large numbers | Probability and Statistics | Khan Academy
11. Central limit theorem | Inferential statistics | Probability and Statistics | Khan Academy
12. Margin of error 1 | Inferential statistics | Probability and Statistics | Khan Academy
13. Margin of error 2 | Inferential statistics | Probability and Statistics | Khan Academy

14. Hypothesis testing and p-values | Inferential statistics | Probability and Statistics | Khan Academy
15. One-tailed and two-tailed tests | Inferential statistics | Probability and Statistics | Khan Academy
16. Type 1 errors | Inferential statistics | Probability and Statistics | Khan Academy
17. Large sample proportion hypothesis testing | Probability and Statistics | Khan Academy
18. Boosting and Bagging

Statistics Post-Quiz

1. Data Science Probability Statistics 17

Product And Experiment Designs



Product sense is an important skill for data scientists. Knowing what to measure on new products and why can help determine whether a product is doing well or not. The funny thing is, sometimes metrics going the way you want them to might not always be good. Sometimes the reason people are spending more time on your website is because webpages might be taking longer to load or other similar problems. This is why metrics are tricky and what you measure is important.

Product And Experiment Design Concepts

1. User Engagement Metrics

2. Data Scientist's Toolbox: Experimental Design -Video
3. A/B Testing Guide
4. 6 Themes Of Metrics

Product And Metrics Questions

1. An important metric goes down, how would you dig into the causes?\
2. What metrics would you use to quantify the success of youtube ads (this could also be extended to other products like Snapchat filters, twitter live-streaming, fort-nite new features, etc)
3. How do you measure the success or failure of a product/product feature
4. Google has released a new version of their search algorithm, for which they used A/B testing. During the testing process, engineers realized that the new algorithm was not implemented correctly and returned less relevant results. Two things happened during testing:

- People in the treatment group performed more queries than the control group.
- Advertising revenue was higher in the treatment group as well.

What may be the cause of people in the treatment group performing more searches than the control group? There are different possible answers here.

Question 4 borrowed from Zarantech; We really enjoyed it and thought it was a good example of how things can go wrong.

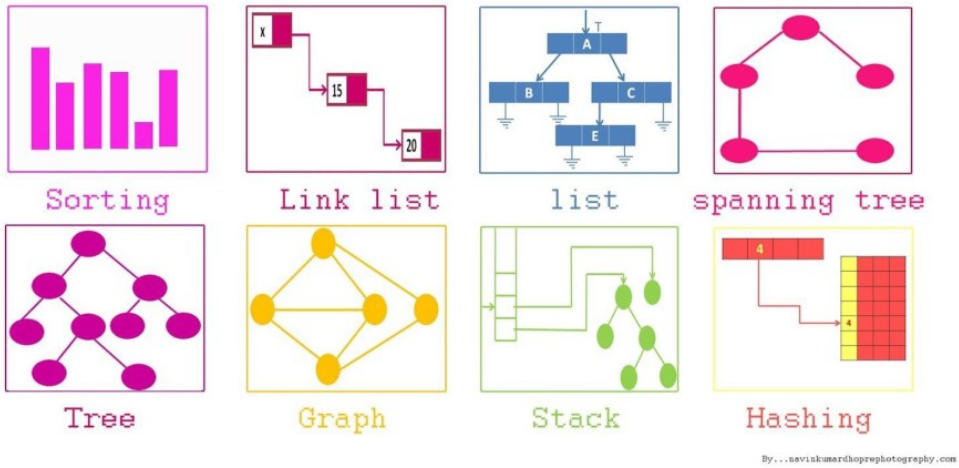
Programming

Just because data science doesn't always require heavy programming, it doesn't mean that interviewers won't ask you traverse a binary tree. So make sure you ask your interviewer what to expect. Don't be daunted by these questions. Pick a few to do just so you're not surprised in an interview.

Pre-Video Questions

1. Fizz Buzz
2. Find The Kth Smallest/Largest Integer In An Array
3. Nth Fibonacci

Algorithms And Data Structures



Pre-Study ProblemsBefore going through the video content about data structures and algorithms. Consider trying out these problems below. See if you can answer them. This will help you know what to focus on.

1. 985. Sum of Even Numbers After Queries
2. 657. Robot Return to Origin
3. 961. N-Repeated Element in Size 2N Array
4. 110. Balanced Binary Tree

Algorithms And Data Structures Videos

Data Structures

1. Data Structures & Algorithms #1—What Are Data Structures?
2. Multi-dim (video)
3. Data Structures: Linked Lists
4. Core Linked Lists Vs Arrays (video)
5. Data Structures: Trees
6. Data Structures: Heaps
7. Data Structures: Hash Tables
8. Data Structures: Stacks and Queues

Algorithms

1. Python Algorithms for Interviews
2. Algorithms: Graph Search, DFS and BFS
3. BFS(breadth-first search) and DFS(depth-first search) (video)
4. Algorithms: Binary Search
5. Binary Search Tree Review (video)
6. Algorithms: Recursion
7. Algorithms: Bubble Sort
8. Algorithms: Merge Sort
9. Algorithms: Quicksort

String Manipulation

1. Coding Interview Question and Answer: Longest Consecutive Characters
2. Sedgewick—Substring Search (videos)

SQL

Post-Study Problems

Now that you have studied for a bit, and watched a few videos. Let's try some more problems!

1. Bigger Is Greater
2. 6. ZigZag Conversion
3. 7. Reverse Integer
4. 40. Combination Sum II
5. 43. Multiply Strings
6. Larry's Array
7. hort Palindrome
8. 65. Valid Number
9. Bigger is Greater
10. The Full Counting Sort
11. Lily's Homework

SQL—Problems

Generally, there will be at least one interview focused on SQL. In addition, the interviewers might take you through an entire process of developing a product, a metric and then a query to measure the effectiveness of that metric.

1. 262. Trips and Users
2. 601. Human Traffic of Stadium
3. 185. Department Top Three Salaries
4. 626. Exchange Seats
5. Hackerrank The Report
6. 177. Nth Highest Salary
7. Symmetric Pairs
8. Occupations
9. Placements
10. Ollivander's Inventory

SQL—Videos

1. IQ15: 6 SQL Query Interview Questions
2. Learning about ROW_NUMBER and Analytic Functions
3. Advanced Implementation Of Analytic Functions
4. Advanced Implementation Of Analytic Functions Part 2

5. Wise Owl SQL Videos

Post SQL Problems

1. Binary Tree Nodes
2. Weather Observation Station 18
3. Challenges
4. Print Prime Numbers
5. 595. Big Countries
6. 626. Exchange Seats
7. SQL Interview Questions: 3 Tech Screening Exercises (For Data Analysts)

Technical interviews can be tough. Whether they are for software engineers, data engineers or data scientists. We do hope this study guide helps you keep track of your progress!

Let us know if you think a subject is missing. Thank you!

If you enjoyed this video about software engineering then consider these videos as well!

The Interview Study Guide For Software Engineers

Learning Data Science: Our Top 25 Data Science Courses

The Best And Only Python Tutorial You Will Ever Need To Watch

Dynamically Bulk Inserting CSV Data Into A SQL Server

4 Must Have Skills For Data Scientists

Engineering Dashboards, Metrics And Algorithms Part 2

Read Last Weeks Top Ten Article For Python Libraries

How Algorithms Can Become Unethical and Biased

142 Resources for Mastering Coding Interviews

CATEGORIES DATA SCIENCE

SHARE THIS ARTICLE

OLDER ARTICLE

10 AMAZING ARTICLES ON PYTHON PROGRAMMING AND MACHINE LEARNING WEEK 4

by Ben R. - May 19, 2019