

This site uses cookies from Google to deliver its services, to personalise ads and to analyse traffic. Information about your use of this site is shared with Google. By using this site, you agree to its use of cookies.

[LEARN MORE](#) [GOT IT](#)



Where Knowledge meets Sharing

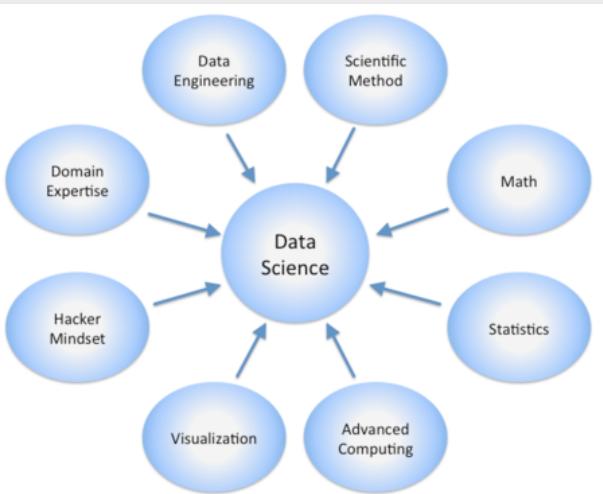
19 October 2015

Data Science Interview Questions



Posted by [Alexey Grigorev](#)

Data Science Interview Questions



Source: [Data Science: An Introduction](#)

Our [IT4BI Master studies](#) finished, and the next logical step after graduation is finding a job. I was interested in Data Science jobs and this post is a summary of my interview experience and preparation.

The term "Data Science" is not yet well establish, so interviews for Data Science jobs might include a very broad range of questions, depending on the interpretation of the term by a particular company. In this post I attempt to organize Data Science interview questions in some usable form, but it might also be biased by how I see Data Science myself. I hope you also can find it useful.

The sources of the questions are:

- links that I discovered on the Internet,
- my own data science interviews (being on the interviewee side)

The questions are without answers. First of all, the answer that I would write could be bad or wrong, and second, the post would be too big. Also, going through the list and looking for the answers yourself is a good exercise to prepare for an interview.

This list might look scary at first, but it's very unlikely that all of these questions will be asked during one interview. Very few jobs require applicants to know all of these points. So it's rather a broad overview of things that may potentially be asked. Don't let this list of questions discourage you if you don't know the answer to some of them: chances are that these questions are not important for your interview.

So, let's get started.

Table of Content

- [Background Questions](#)

Our Posts

- ▼ [2015 \(18\)](#)
- ▼ [October \(3\)](#)
 - [Data Science Interview Questions](#)
 - [Java Interview Questions](#)
 - [Mastering Data Analysis with R](#)
- [June \(3\)](#)
- [April \(2\)](#)
- [March \(3\)](#)
- [February \(2\)](#)
- [January \(5\)](#)
- [2014 \(2\)](#)

Who Are We?

We are a group of Master's Degree students from all over the world who met each other thanks to the foundation and specialisation courses of the *Erasmus Mundus IT4BI Programme*.



Ahmet Anil Pala



Alexey Grigorev



Andres Felipe Zamora Montaño



Elena Samota



Guven Toprakkiran



Hicham Akaoka Badssi



José Luis Pino López



Madalina Burghelea



Maximiliano Ariel López



This site uses cookies from Google to deliver its services, to personalise ads and to analyse traffic. Information about your use of this site is shared with Google. By using this site, you agree to its use of cookies.

[LEARN MORE](#) [GOT IT](#)

- Linear Algebra
- Other Areas
- Probability and Statistics
 - Basic Probability
 - Distributions
 - Basic Statistics
 - Experiment Design
 - Point Estimates
 - Testing
 - A/B Tests
 - Bayesian Statistics
 - Time Series
 - Advanced
- Machine Learning
 - General ML Questions
 - Regression
 - Classification
 - Regularization
 - Dimensionality Reduction
 - Cluster Analysis
 - Optimization
 - Recommendation
 - Feature Engineering
 - Natural Language Processing
 - Meta Learning
 - Miscellanea
- Computer Science
 - Libraries and Tools
 - Databases
 - Distributed Systems and Big Data
- Hands-On
 - Problem to Solve
 - Coding
 - Papers
- Sources
- Useful Links

Background Questions

Usually, interviews start with background questions: they can ask you to talk about yourself. This can also happen at the telephone interview stage.

Careers

For background questions be ready to talk about a summary of your career.

- Summarize your experience
- What companies you worked at? What was your role?
- Do you have a project portfolio? What projects you implemented? Discuss some of them in details
- For graduating students: Tell me about your master thesis
- For aspiring data scientists: Why do you want a career in data science?
- What are your career goals?

General Questions



Navid Mahlouji



Nyami Ronald
Mitterand



Steffi Melinda



Stephany García
Martínez



Tamara Mendt



This site uses cookies from Google to deliver its services, to personalise ads and to analyse traffic. Information about your use of this site is shared with Google. By using this site, you agree to its use of cookies.

[LEARN MORE](#) [GOT IT](#)

- What data science blogs do you follow?
- Have you taken any data-science-related online courses? If yes, how many did you complete with a certificate?

Process

All Machine Learning, Data Mining and Data Science projects should follow some process, so there can be questions about it:

- Can you outline the steps in an analytics project?
- Have you heard of CRISP-DM (Cross Industry Standard Process for Data Mining)?

CRISP-DM defines the following steps:

- Problem Definition
- Data Understanding (or Data Exploration)
- Data Preparation
- Modeling
- Evaluation
- Deployment (for the production)

So next you may discuss each of these steps in details

- What is the goal of each step?
- What are possible activities at each step?

Mathematics

Some background mathematics is necessary for doing Data Science, therefore you should expect math-related questions.

Linear Algebra

Basic Linear Algebra questions might include:

- What is $Ax = b$? How to solve it?
- How do we multiply matrices?
- What is an Eigenvalue? And what is an Eigenvector? What is Eigenvalue Decomposition or The Spectral Theorem?
- What is Singular Value Decomposition?
- You can expect tons of Liner Algebra questions in the Machine Learning part of the interview (see below).

If you are interested in learning or refreshing Linear Algebra, see [Best Time to Learn Linear Algebra is Now!](#)

Other Areas

- Discrete Mathematics and Logics are not that important for Data Science
- Probability and Statistics are core skills and discussed in the next section
- Calculus and Optimization are usually discussed in the Machine Learning part and usually when talking about a particular algorithm

Probability and Statistics

Probability and Statistics is an important part of an interview, because it's the basics for Machine Learning. It is also useful if the company is doing some marketing or website optimization, so they could ask about related concepts such as A/B tests.

This site uses cookies from Google to deliver its services, to personalise ads and to analyse traffic. Information about your use of this site is shared with Google. By using this site, you agree to its use of cookies.

[LEARN MORE](#) [GOT IT](#)

- Given two fair dices, what is the probability of getting scores that sum to 4? to 8?
- A simple questions on Bayes rule: Imagine a test with a true positive rate of 100% and false positive rate of 5%. Imagine a population with a 1/1000 rate of having the condition the test identifies. Given a positive test, what is the probability of having that condition?

Distributions

You can expect questions about probability distributions:

- What is the normal distribution? Give an example of some variable that follows this distribution
- What about log-normal?
- Explain what a long tailed distribution is and provide three examples of relevant phenomena that have long tails. Why are they important in classification and prediction problems?
- How to check if a distribution is close to Normal? Why would you want to check it? What is a QQ Plot?
- Give examples of data that does not have a Gaussian distribution, or log-normal.
- Do you know what the exponential family is?
- Do you know the Dirichlet distribution? the multinomial distribution?

Basic Statistics

- What is the Laws of Large Numbers? Central Limit Theorem?
- Why are they important for Statistics?
- What summary statistics do you know?

Experiment Design

Designing experiments is an important part of Statistics, and it's especially useful for doing A/B tests.

Sampling and Randomization

- Why do we need to sample and how?
- Why is randomization important in experimental design?
- Some 3rd party organization randomly assigned people to control and experiment groups. How can you verify that the assignment truly was random?
- How do you calculate needed sample size?
- Power analysis. What is it?

Biases

- When you sample, what bias are you inflicting?
- How do you control for biases?
- What are some of the first things that come to mind when I do X in terms of biasing your data?

Other questions

- What are confounding variables?

Point Estimates

Confidence intervals

- What is a point estimate? What is a confidence interval for it?
- How are they constructed?
- Why do you need to standardize?
- How to interpret confidence intervals?

Testing

Hypothesis tests

This site uses cookies from Google to deliver its services, to personalise ads and to analyse traffic. Information about your use of this site is shared with Google. By using this site, you agree to its use of cookies.

[LEARN MORE](#) [GOT IT](#)

- How would you test if two populations have the same mean? What if you have 3 or 4 populations?
- You applied ANOVA and it says that the mean is different. How do you identify the populations where the means are different?
- What is the distribution of p-value's, in general?

A/B Tests

- What is A/B testing? How is it different from usual Hypothesis testing?
- How can you prove that one improvement you've brought to an algorithm is really an improvement over not doing anything? How familiar are you with A/B testing?
- How can we tell whether our website is improving?
- What are the metrics to evaluate a website? A search engine?
- What kind of metrics would you track for your music streaming website?
- Common metrics: Engagement / retention rate, conversion, similar products / duplicates matching, how to measure them.
- Real-life numbers and intuition: Expected user behavior, reasonable ranges for user signup / retention rate, session length / count, registered / unregistered users, deep / top-level engagement, spam rate, complaint rate, ads efficiency.

Bayesian Statistics

In my interviews I didn't have any questions about Bayesian Stats, nor did I find a lot of questions on the Internet. But here are some:

- Have you ever seen Bayes Theorem?
- Do you know what a conjugate-prior is?

You might also get questions about Bayesian non-parametric models, but I'm not sure if it's common.

Time Series

- What is a time series?
- What is the difference between data for usual statistical analysis and time series data?
- Have you used any of the following: Time series models, Cross-correlations with time lags, Correlograms, Spectral analysis, Signal processing and filtering techniques? If yes, in which context?
- In time series modeling how can we deal with multiple types of seasonality like weekly and yearly seasonality?

Advanced

Resampling

- Explain what resampling methods are and why they are useful. Also explain their limitations.
- Bootstrapping - how and why is it used?
- How to use resampling for hypothesis testing? Have you heard of Permutation Tests?
- How would you apply resampling to time series data?

Machine Learning

In my experience, the Machine Learning part is usually the largest part of the interview. It may be a few basic questions, but it's helpful to be prepared to more in-depth Machine Learning questions, especially if you claim to have worked with it on your CV.

General ML Questions

The ML part may start with something like:

- What is the difference between supervised and unsupervised learning? Which algorithms are supervised learning and which are not? Why?

This site uses cookies from Google to deliver its services, to personalise ads and to analyse traffic. Information about your use of this site is shared with Google. By using this site, you agree to its use of cookies.

[LEARN MORE](#) [GOT IT](#)

Regression

- Describe the regression problem. Is it supervised learning? Why?
- What is linear regression? Why is it called linear?
- Discuss the bias-variance tradeoff.

Linear Regression:

- What is Ordinary Least Squares Regression? How it can be learned?
- Can you derive the OLS Regression formula? (For one-step solution)
- Is model $Y = X_1 + X_2 + X_1 X_2 + \epsilon$ still linear? Why?
- Do we always need the intercept term? When do we need it and when do we not?
- What is collinearity and what to do with it? How to remove multicollinearity?
- What if the design matrix is not full rank?
- What is overfitting a regression model? What are ways to avoid it?
- What is Ridge Regression? How is it different from OLS Regression? Why do we need it?
- What is Lasso regression? How is it different from OLS and Ridge?

Linear Regression assumptions:

- What are the assumptions required for linear regression?
- What if some of these assumptions are violated?

Significant features in Regression

- You would like to find significant features. How would you do that?
- You fit a multiple regression to examine the effect of a particular feature. The feature comes back insignificant, but you believe it is significant. Why can it happen?
- Your model considers the feature X significant, and Z is not, but you expected the opposite result. Why can it happen?

Evaluation

- How to check is the regression model fits the data well?

Other algorithms for regression

- Decision trees for regression
- k -Nearest Neighbors for regression. When to use?
- Do you know others? E.g. Splines? LOESS/LOWESS?

Classification

Basic:

- Can you describe what is the classification problem?
- What is the simplest classification algorithm?
- What classification algorithms do you know? Which one you like the most?

Decision trees:

- What is a decision tree?
- What are some business reasons you might want to use a decision tree model?
- How do you build it?
- What impurity measures do you know?
- Describe some of the different splitting rules used by different decision tree algorithms.
- Is a big bushy tree always good? Why would you want to prune it?
- Is it a good idea to combine multiple trees?
- What is Random Forest? Why is it good?

Logistic regression:

- What is logistic regression?
- How do we train a logistic regression model?
- How do we interpret its coefficients?

Support Vector Machines

- What is the maximal margin classifier? How this margin can be achieved and why is it

This site uses cookies from Google to deliver its services, to personalise ads and to analyse traffic. Information about your use of this site is shared with Google. By using this site, you agree to its use of cookies.

[LEARN MORE](#) [GOT IT](#)

Neural Networks

- What is an Artificial Neural Network?
- How to train an ANN? What is back propagation?
- How does a neural network with three layers (one input layer, one inner layer and one output layer) compare to a logistic regression?
- What is deep learning? What is CNN (Convolution Neural Network) or RNN (Recurrent Neural Network)?

Other models:

- What other models do you know?
- How can we use Naive Bayes classifier for categorical features? What if some features are numerical?
- Tradeoffs between different types of classification models. How to choose the best one?
- Compare logistic regression with decision trees and neural networks.

Regularization

- What is Regularization?
- Which problem does Regularization try to solve?
- What does it mean (practically) for a design matrix to be "ill-conditioned"?
- When might you want to use ridge regression instead of traditional linear regression?
- What is the difference between the L_1 and L_2 regularization?
- Why (geometrically) does LASSO produce solutions with zero-valued coefficients (as opposed to ridge)?
- Let us go through the derivation of OLS or Logistic Regression. What happens when we add L_2 regularization? How do the derivations change? What if we replace L_2 regularization with L_1 regularization?

Dimensionality Reduction

Basics:

- What is the purpose of dimensionality reduction and why do we need it?
- Are dimensionality reduction techniques supervised or not? Are all of them are (un)supervised?
- What ways of reducing dimensionality do you know?
- Is feature selection a dimensionality reduction technique?
- What is the difference between feature selection and feature extraction?
- Is it beneficial to perform dimensionality reduction before fitting an SVM? Why or why not?

Principal Component Analysis:

- What is Principal Component Analysis (PCA)? What is the problem it solves? How is it related to eigenvalue decomposition (EVD)?
- What's the relationship between PCA and SVD? When SVD is better than EVD for PCA?
- Under what conditions is PCA effective?
- Why do we need to center data for PCA and what can happen if we don't do it? Do we need to scale data for PCA?
- Is PCA a linear model or not? Why?

Other Dimensionality Reduction techniques:

- Do you know other Dimensionality Reduction techniques?
- What is Independent Component Analysis (ICA)? What's the difference between ICA and PCA?
- Suppose you have a very sparse matrix where rows are highly dimensional. You project these rows on a random vector of relatively small dimensionality. Is it a valid dimensionality reduction technique or not?
- Have you heard of Kernel PCA or other non-linear dimensionality reduction techniques? What about LLE (Locally Linear Embedding) or t-SNE (t -distributed Stochastic Neighbor Embedding)
- What is Fisher Discriminant Analysis? How is it different from PCA? Is it supervised or not?

This site uses cookies from Google to deliver its services, to personalise ads and to analyse traffic. Information about your use of this site is shared with Google. By using this site, you agree to its use of cookies.

[LEARN MORE](#) [GOT IT](#)

- Describe K -Means. What is the objective of K -Means? Can you describe the Lloyd algorithm?
- How do you select K for K-Means?
- How can you modify K -Means to produce soft class assignments?
- How to assess the quality of clustering?
- Describe any other cluster analysis method. E.g. DBSCAN.

Optimization

You may have some basic questions about optimization:

- What is the difference between a convex function and non-convex?
- What is Gradient Descent Method?
- Will Gradient Descent methods always converge to the same point?
- What is a local optimum?
- Is it always bad to have local optima?
- What is Newton's method?
- What kind of problems are well suited for Newton's method? BFGS? SGD?
- What are "slack variables"?
- Describe a constrained optimization problem and how you would tackle it.

Recommendation

- What is a recommendation engine? How does it work?
- Do you know about the Netflix Prize problem? How would you approach it?
- How to do customer recommendation?
- What is Collaborative Filtering?
- How would you generate related searches for a search engine?
- How would you suggest followers on Twitter?

Feature Engineering

- How to apply Machine Learning to audio data, images, texts, graphs, etc?
- What is Feature Engineering? Can you give an example? Why do we need it?
- How to go from categorical variables to numerical?

Natural Language Processing

If the company deals with text data, you can expect some questions on NLP and Information Retrieval:

- What is NLP? How is it related to Machine Learning?
- How would you turn unstructured text data into structured data usable for ML models?
- What is the Vector Space Model?
- What is TF-IDF?
- Which distances and similarity measures can we use to compare documents? What is cosine similarity?
- Why do we remove stop words? When do we not remove them?
- Language Models. What is N -Grams?

Meta Learning

Feature Selection:

- Are all features equally good?
- What are the downfalls of using too many or too few variables?
- How many features should you use? How do you select the best features?
- What is Feature Selection and why do we need it?
- Describe several feature selection methods. Are these methods depend on the model or

This site uses cookies from Google to deliver its services, to personalise ads and to analyse traffic. Information about your use of this site is shared with Google. By using this site, you agree to its use of cookies.

[LEARN MORE](#) [GOT IT](#)

- You have one model and want to find the best set of parameters for this model. How would you do that?
- How would you look for the best parameters? Do you know something else apart from grid search?
- What is Cross-Validation?
- What is 10-Fold CV?
- What is the difference between holding out a validation set and doing 10-Fold CV?

Model evaluation

- How do you know if your model overfits?
- How do you assess the results of a logistic regression?
- Which evaluation metrics you know? Something apart from accuracy?
- Which is better: Too many false positives or too many false negatives?
- What precision and recall are?
- What is a ROC curve? What is AU ROC (AUC)? How to interpret the curve and AU ROC?
- Do you know about Concordance or Lift?

Discussion Questions:

- You have a marketing campaign and you want to send emails to users. You developed a model for predicting if a user will reply or not. How can you evaluate this model? Is there a chart you can use?

Miscellanea

Curse of Dimensionality

- What is Curse of Dimensionality? How does it affect distance and similarity measures?
- What are the problems of large feature space? How does it affect different models, e.g. OLS? What about computational complexity?
- What dimensionality reductions can be used for preprocessing the data?
- What is the difference between density-sparse data and dimensionally-sparse data?

Others

- You are training an image classifier with limited data. What are some ways you can augment your dataset?

Computer Science

Knowledge in Computer Science is as important for Data Science as knowledge in Machine Learning. So you may get the same type of questions as for any software developer position, but possibly with lower expectations on your answers.

I was a Java developer for quite some time, and I prepared a list of questions I asked (and often was asked) on Java interviews: [Java Interview questions](#). This list can also be helpful for preparing to a Data Science interview.

Libraries and Tools

Apart from basics of Java/Scala/Python/etc, you may be asked about libraries for data analysis:

- Which libraries for data analysis do you know in Python/R/Java?
- Have you used numpy, scipy, pandas, sklearn?
- What are some features of the sklearn api that differentiate it from fitting models in R?
- What are some features of pandas/scipy that you like? Hate? Same questions for R.
- Why is “vectorization” such a powerful method for optimizing numerical code? What is going on that makes the code faster relative to alternatives like nested for loops?
- When is it better to write your own code than using a data science software package?
- State any 3 positive and negative aspects about your favorite statistical software.
- Describe a difficult bug you've encountered and how you resolved it.
- How does floating point affect precision of calculations? Equality tests?
- What is BLAS? LAPACK?

This site uses cookies from Google to deliver its services, to personalise ads and to analyse traffic. Information about your use of this site is shared with Google. By using this site, you agree to its use of cookies.

[LEARN MORE](#) [GOT IT](#)

- Or given some DB schema you may be asked to write a simple SQL query.
- What is a "star schema"? "snowflake schema"?
- Describe different NoSQL technologies you're familiar with, what they are good at, and what they are bad at.

Distributed Systems and Big Data

Basic "Big Data" questions:

- What is the biggest data set that you have processed and how did you process it? What was the result?
- Have you used Apache Hadoop, Apache Spark, Apache Flink? Why? Have you used Apache Mahout?

MapReduce

- What are the advantages/disadvantages of "shared-nothing" architecture?
- What is MapReduce? Why is it "shared-nothing" architecture?
- Can you implement word count in MapReduce? What about something a bit more complex like TF-IDF? Naive Bayes?
- What is load balance? How to make sure a MapReduce application has good load balance?
- Can you give examples where MapReduce does not work?
- What are examples of "embarrassingly parallelizable" algorithms?

Implementation questions

- How would you estimate the median of a dataset that is too big to hold in the memory?

There are some posts that you may find useful when preparing for the "Big Data" part:

- [Hadoop and MapReduce](#)
- [Naive Bayes on Apache Flink](#)

Hands-On

Also, many interviews have a part which I call "hands-on": you are given some problem description and you are asked to solve it. You can just talk the interviewers through your solution or even be asked to sit and implement some parts. Sometimes there is also a test assignment to be done at home (prior to the interview).

Problem to Solve

For example:

Assume that you are asked to lead a project on churn detection, and have dataset of known users who stopped using the service and ones who are still using. This data includes demographics and other features.

Do the following:

1. Describe the methodology and model that you will chose to identify churn, and describe your thought process.
2. Think how would you communicate the results to the CEO?
3. Suppose in the dataset only 0.025 of users churned. How would you make it more balanced?

Also:

- How would you implement it if you had one day? One month? One year?
- How would your approach scale?

Other problems:

- How would you approach identifying plagiarism?
- How to find individual paid accounts shared by multiple users?

This site uses cookies from Google to deliver its services, to personalise ads and to analyse traffic. Information about your use of this site is shared with Google. By using this site, you agree to its use of cookies.

[LEARN MORE](#) [GOT IT](#)

- How would you approach collecting the data if you didn't have the dataset?

Coding

Sometimes you even may be presented a small dataset and ask to do a particular task with any tool. For example,

- write a script to extract features,
- then do some exploratory data analysis and
- finally apply some ML algorithm to this dataset.

Or just the last two, with a ready to use dataset in tabular form.

Papers

It's also possible that you'll be asked to read some ML paper and share your thoughts on it, and then discuss the proposed algorithm, its time complexity, how it can be implemented and improved.

I wasn't asked to do it myself, but based on my experience working as a ML developer, I believe that reading papers and being able to understand them is an important skill, so don't be surprised if somebody tries to check this ability.

Sources

I had to work through a lot of sources to make this compilation. I did not include all the questions I came across, just the ones that made sense or ones I really got during my interviews. It also, of course, includes my own interviews.

Here is the list of sources I used:

- <http://www.quora.com/What-is-a-typical-data-scientist-interview-like>
- <http://www.quora.com/What-are-the-interview-questions-on-regression-modeling>
- <http://www.quora.com/How-should-I-prepare-for-statistics-questions-for-a-data-science-interview>
- <http://www.quora.com/A-B-Testing/What-kind-of-A-B-testing-questions-should-I-expect-in-a-data-scientist-interview-and-how-should-I-prepare-for-such-questions>
- <http://www.quora.com/What-are-20-questions-to-detect-fake-data-scientists>
- <http://www.quora.com/What-are-some-common-Machine-Learning-interview-questions>
- <http://www.quora.com/What-are-the-best-interview-questions-to-evaluate-a-machine-learning-researcher>
- <https://www.quora.com/Are-CS-questions-part-of-a-data-scientist-interview-at-Facebook>
- <http://www.quora.com/Data-Science/How-should-I-prepare-for-statistics-questions-for-a-data-science-interview>
- <http://stats.stackexchange.com/questions/5465/statistics-interview-questions>
- http://www.reddit.com/r/datascience/comments/2nhb4k/what_interview_questions_have_you_been_asked/
- http://www.reddit.com/r/statistics/comments/310h76/i_have_an_interview_for_a_parttime_data_analyst/
- https://www.reddit.com/r/datascience/comments/3fsz54/my_top10_technical_questions_for_job_candidates/
- https://www.reddit.com/r/datascience/comments/3kzf69/data_scientist_interview_questions_on_pca_svm/
- http://www.reddit.com/r/MachineLearning/comments/392nwq/interview_questions_for_data_scientist_positions/
- <http://blog.udacity.com/2015/04/data-science-interview-questions.html>
- <http://alyaababb.wordpress.com/2014/10/01/how-to-ace-a-data-science-interview/>
- <http://www.marketingdistillery.com/2014/09/03/how-to-successfully-recruit-a-data-scientist/>
- <http://www.edureka.co/blog/frequently-asked-data-science-interview-questions>
- <http://www.galvanize.it/blog/how-to-nail-a-data-science-interview>
- <http://analyticsindiamag.com/common-analytics-interview-questions/>

This site uses cookies from Google to deliver its services, to personalise ads and to analyse traffic. Information about your use of this site is shared with Google. By using this site, you agree to its use of cookies.

[LEARN MORE](#) [GOT IT](#)

Useful Links

If you are preparing to a Data Science interview, you may also find the following links useful:

- <http://www2.udacity.com/rs/udacity/images/Ultimate%20Skills%20Checklist%20For%20Your%20First%20Data%20Analyst%20Job.pdf>
- <http://www.quora.com/What-are-some-important-questions-to-ask-a-recruiter-when-interviewing-for-a-data-science-job>
- <http://www.quora.com/In-a-data-scientist-interview-should-I-use-Python-or-C++-for-algorithm-data-structure-questions>
- <http://www.quora.com/How-do-I-prepare-for-a-data-scientist-interview>
- <http://datascienceinterview.quora.com/Data-Science-Interview-Preparation>
- <http://datascienceinterview.quora.com/Answers-1>
- <https://github.com/gkamradt/Lessons-Learned-Data-Science-Interviews>
- <http://matthewanalytics.com/2015/08/18/homework-during-the-hiring-process-no-thanks/>
- <https://medium.com/@D33B/interview-questions-for-data-scientist-positions-5ad3c5d5b8bd>
- <https://medium.com/@D33B/interview-questions-for-data-scientist-positions-part-ii-ac294c2c7241>
- <http://www.jasq.org/just-another-scala-quant/new-agey-interviews-at-the-grocery-startup>
- <http://www.erinhellman.com/crushed-it-landing-a-data-science-job/>
- http://treycausey.com/data_science_interviews.html

The End

Even though the post was lengthy, I hope you enjoyed it and found this information useful. Happy interviewing! And please do let us know if you got any interesting questions that we should add.



+11 Recommend this on Google

Labels: by Alexey, Data Science, Interview Questions, Machine Learning

3 comments:



ming 20/10/15 23:20

Thank you for your sharing!

[Reply](#)



Dave Od 23/10/15 14:17

fantastic overall view of data science and associated technologies and methods - thanks

[Reply](#)



Unknown 25/10/15 16:09

amazing content, thanks a bunch!

[Reply](#)

This site uses cookies from Google to deliver its services, to personalise ads and to analyse traffic. Information about your use of this site is shared with Google. By using this site, you agree to its use of cookies.

[LEARN MORE](#) [GOT IT](#)

Comment as: CheeseBiscuit (Gc)

[Sign out](#)

[Publish](#)

[Preview](#)

[Notify me](#)

[Home](#)

[Older Post](#)

Subscribe to: [Post Comments \(Atom\)](#)

Powered by [Blogger](#).