

Home (https://www.educba.com/) » Blog (https://www.educba.com/blog/) » Data and Analytics (https://www.educba.com/category/analytics-blog/) » Top 5 Most Valuable Data Science Interview Questions

Unlimited access to all 551 Courses & 204 Job Oriented Programs

\$99

FOR ANNUAL

~~\$199~~

Expires in

00 D 05 H 46 M 46 S

ENROLL NOW

(HTTPS://WWW.EDUCBA.COM/SHOPPING_CART.PHP?PID=5335&PTYPE=999)

enter email id

Give me FREE Access

Job Oriented Programs

Data Science All in One Bundle (https://www.educba.com/job-oriented-programs/data-science)

Become a Hadoop Developer (https://www.educba.com/job-oriented-programs/become-a-hadoop-developer)

Become a SAS Developer (https://www.educba.com/job-oriented-programs/become-a-sas-developer)

Become a Machine Learning Scientist (https://www.educba.com/job-oriented-programs/become-a-machine-learning-scientist)

Become an R Programmer (https://www.educba.com/job-oriented-programs/become-an-r-programmer)

View All

(https://www.educba.com/job-oriented-programs/all/data-science/)

^

educba.com/library/

Price: \$99

https://www.educba.com/data-science-interview-questions/

1/5



Introduction to Data Science interview questions and answers

If you are looking for a job that is related to Data Science, you need to prepare for the Data science interview questions. Though every Data Science interview is different and the scope of a job is also different, we can help you out with the top Data Science interview questions and answers, which will help you take the leap and get you success in your interview.

Below is the list of Data Science Interview Questions that are mostly asked

1.What is Data Science?

Answers: Data Science is an interdisciplinary field of different scientific methods, techniques, processes, and knowledge that is used to transform the data of different types such as **structured** (<https://www.educba.com/what-is-structured-finance/>), unstructured and semi-structured data into the required format or representation.

Data Science concepts include different concepts such as statistics, regression, mathematics, **computer science** (<https://www.educba.com/computer-science-vs-data-science/>), algorithms, **data structures** (<https://www.educba.com/course/data-structures-javascript/>) and information science with also including some subfields such as **data mining** (<https://www.educba.com/data-mining-vs-machine-learning/>), **machine learning** (<https://www.educba.com/big-data-vs-machine-learning/>), and **databases** (<https://www.educba.com/database-management-system/>) etc.,

Data Science concept has recently evolved to a greater extent in the area of computing technology in order to perform **data analysis** (<https://www.educba.com/data-analysis-techniques-for-brand-strength/>) on the existing data where the growth of data is in terms of an exponential with respect to time.

Data Science is the study of various types of data such as structured, semi-structured and unstructured data in any form or formats available in order to get some information out of it.

Data Science consists of different technologies used to study data such as data mining, data storing, data purging, data archival, data transformation etc., in order to make it efficient and ordered. Data Science also includes the concepts like Simulation, modeling, analytics, machine learning, computational mathematics etc.,

2.What is the best Programming Language to use in Data Science?

Answers: Data Science can be handled by using programming languages like Python or R programming language. These two are the most popular languages being used by the Data Scientists (<https://www.educba.com/>) or Data Analysts. R and Python (<https://www.educba.com/python-and-django-for-web-development/>) are open source and are free to use and came into existence during the 1990s. Pricing (<https://www.educba.com/plans/>)

Python and R have different advantages depending on the applications and required a business goal. Python is better to be used in the cases of repeated tasks or jobs and for data manipulations whereas **R programming** (<https://www.educba.com/course/beginners-training-r-programming/>) can be used for querying or retrieving datasets and customized data analysis.

Mostly Python is preferred for all types of data science applications where some time R programming is preferred in the cases of high or complex data applications. Python is easier to learn and has less learning curve whereas R has a deep learning curve.

Python is mostly preferred in all the cases which is a general-purpose programming language and can be found in many applications other than Data Science too. R is mostly seen in Data Science area only where it is used for data analysis in standalone servers or computing separately.

3. Why is data cleaning essential in Data Science?

Answers: Data cleaning is more important in Data Science because the end results or the outcomes of the data analysis come from the existing data where useless or unimportant need to be cleaned periodically as of when not required. This ensures the data reliability & accuracy and also memory is freed up.

Data cleaning reduces the data redundancy and gives good results in data analysis where some large customer information exists and that should be cleaned periodically. In the businesses like e-commerce, retail, government organizations contain large customer transaction information which is outdated and needs to be cleaned.

Depending on the amount or size of data, suitable tools or methods should be used to clean the data from the database or **big data** (<https://www.educba.com/big-data-vs-data-science/>) environment. There are different types of data existing in a data source such as dirty data, clean data, mixed clean and dirty data and sample clean data.

Modern data science applications rely on **machine learning** (<https://www.educba.com/statistics-vs-machine-learning/>) model where the learner learns from the existing data. So, the existing data should always be cleanly and well maintained to get sophisticated and good outcomes during the optimization of the system.

4. What is Linear Regression in Data Science?

Answers: Linear Regression is a technique used in **supervised machine learning** (<https://www.educba.com/supervised-learning-vs-reinforcement-learning/>) algorithmic process in the area of Data Science. This method is used for **predictive analysis** (<https://www.educba.com/predictive-analysis-vs-forecasting/>).

Predictive analytics (<https://www.educba.com/predictive-analytics-vs-data-science/>) is an area within Statistical Sciences where the existing information will be extracted and processed to predict the trends and outcomes pattern. The core of the subject lies in the analysis of existing context to predict an unknown event.

The process of Linear Regression method is to predict a variable called target variable by making the best relationship between the dependent variable and an independent variable. Here dependent variable is outcome variable and also response variable whereas the independent variable is predictor variable or explanatory variable.

For example in real life, depending on the expenses occurred in this financial year or monthly expenses, the predictions happen by calculating the approximate upcoming months or financial years expenses.

In this method, the implementation can be done by using **Python programming** (<https://www.educba.com/course/python-programming-zero-to-hero/>) technique where this is the most important method used in Machine Learning technique under the area of Data Science.

5. What is A/B testing in Data Science?

Answers: A/B testing is also called as Bucket Testing or Split Testing. This is the method of comparing and testing two versions of systems or applications against each other to determine which version of application performs better. This is important in the cases where multiple versions are shown to the customers or end users in order to achieve the goals.

In the area of Data Science (https://www.educba.com/data-science-and-its-growing-importance/), this A/B testing is used to know which variable out of the existing two variables in order to optimize or increase the outcome of the goal. A/B testing is also called Design of Experiment. This testing helps in establishing a cause and effect relationship between the independent and dependent variables.

This testing is also simply a combination of design experimentation or statistical inference. Significance, Randomization and Multiple Comparisons are the key elements of the A/B testing.

The significance is the term for the significance of statistical tests conducted. Randomization is the core component of the experimental design where the variables will be balanced. Multiple comparisons are the way of comparing more variables in the case of customer interests that causes more false positives resulting in the requirement of correction in the confidence level of a seller in the area of e-commerce.

A/B testing is the important one in the area of Data Science in predicting the outcomes.

Recommended Article

This has been a guide to Basic List Of Data Science Interview Questions and answers so that the candidate can crackdown these Data Science Interview Questions easily. You may also look at the following articles to learn more -

1. 5 Effective Interview Grooming Tips For Men (https://www.educba.com/interview-grooming-tips/)
2. Credit Analyst Interview Questions (https://www.educba.com/credit-analyst-interview-questions/)
3. 4 Exciting Interview Preparation Tips To Remember! (https://www.educba.com/interview-preparation-tips/)
4. 10 Excellent MBA Interview Questions You Must Know!!! (https://www.educba.com/mba-interview-questions/)

f	🐦	g
in		

Comments

^

[\(https://www.educba.com/\)](https://www.educba.com/)[\(https://www.educba.com/\)](https://www.educba.com/)[Pricing \(https://www.educba.com/plans/\)](https://www.educba.com/plans/)

Learning

MENU

[Job Oriented Programs \(https://www.educba.com/job-oriented-programs/\)](https://www.educba.com/job-oriented-programs/)[All Courses \(https://www.educba.com/library/\)](https://www.educba.com/library/)[Most Searched Skills \(https://www.educba.com/most-searched-skills/\)](https://www.educba.com/most-searched-skills/)[Blog \(https://www.educba.com/blog/\)](https://www.educba.com/blog/)[Resources \(https://www.educba.com/resources/\)](https://www.educba.com/resources/)[Webinar \(https://www.educba.com/webinar/\)](https://www.educba.com/webinar/)

About Us

[Who is EDUCBA? \(https://www.educba.com/about-us/\)](https://www.educba.com/about-us/)[Careers \(https://www.educba.com/careers/\)](https://www.educba.com/careers/)[Become an Instructor \(https://www.educba.com/instructors/\)](https://www.educba.com/instructors/)[How does it work? \(https://www.educba.com/how-it-works/\)](https://www.educba.com/how-it-works/)[Referral Program \(https://www.educba.com/partners/\)](https://www.educba.com/partners/)[Reseller \(https://www.educba.com/reseller/\)](https://www.educba.com/reseller/)

Support

[Contact Us \(https://www.educba.com/contact-us/\)](https://www.educba.com/contact-us/)[FAQ \(https://www.educba.com/FAQ/\)](https://www.educba.com/FAQ/)[Reviews \(https://www.educba.com/reviews/\)](https://www.educba.com/reviews/)[Register with Redemption Code \(https://www.educba.com/register-with-educba.php\)](https://www.educba.com/register-with-educba.php)[Verifiable Certificate \(https://www.educba.com/verifiable-certificate/\)](https://www.educba.com/verifiable-certificate/)

Apps

[iPhone & iPad \(https://itunes.apple.com/in/app/educba-learning-app/id1341654580?mt=8\)](https://itunes.apple.com/in/app/educba-learning-app/id1341654580?mt=8)[Android \(https://play.google.com/store/apps/details?id=com.educba.www\)](https://play.google.com/store/apps/details?id=com.educba.www)[Terms & Conditions \(https://www.educba.com/terms-and-conditions/\)](https://www.educba.com/terms-and-conditions/)[Disclaimer \(https://www.educba.com/disclaimer/\)](https://www.educba.com/disclaimer/)[Privacy Policy & Cookie Policy \(https://www.educba.com/privacy-policy/\)](https://www.educba.com/privacy-policy/)[Shipping Policy \(https://www.educba.com/shipping-policy/\)](https://www.educba.com/shipping-policy/) [\(https://www.facebook.com/CorporateBridgeGroup\)](https://www.facebook.com/CorporateBridgeGroup) [\(https://plus.google.com/u/0/106267634931284137919/\)](https://plus.google.com/u/0/106267634931284137919/) [\(https://twitter.com/corporatebridge\)](https://twitter.com/corporatebridge) [\(https://www.linkedin.com/company/946029\)](https://www.linkedin.com/company/946029)