

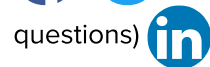
# Data Science Interview Questions

Posted in Data Science (<https://hackr.io/blog/category/data-science>), Interview Questions (<https://hackr.io/blog/category/interview-questions>)



**Akhil Bhadwal**<https://hackr.io/blog/author/akhilbhadwal>

Last Updated 29 Apr, 2020

**Share:**[https://twitter.com/intent/tweet?text=Data+Science+Interview+Questions+https%3A%2F%2Fhackr.io%2Fblog%2Fdata-science-interview-](https://twitter.com/intent/tweet?text=Data+Science+Interview+Questions+https%3A%2F%2Fhackr.io%2Fblog%2Fdata-science-interview-questions)<https://www.linkedin.com/shareArticle?mini=true&url=https://hackr.io/blog/data-science-interview-questions><http://www.reddit.com/submit?url=https://hackr.io/blog/data-science-interview-questions>[https://news.ycombinator.com/submitlink?](https://news.ycombinator.com/submitlink?u=https://hackr.io/blog/data-science-interview-questions)[u=https://hackr.io/blog/data-science-interview-questions](https://hackr.io/blog/data-science-interview-questions)[https://api.whatsapp.com/send?text=https%3A%2F%2Fhackr.io%2Fblog%2Fdata-](https://api.whatsapp.com/send?text=https%3A%2F%2Fhackr.io%2Fblog%2Fdata-science-interview-questions)[science-interview-questions](https://hackr.io/blog/data-science-interview-questions))

1 Comment

## Table of Contents



Data Science is getting bigger and better with each passing day. As such, it is churning out plenty of opportunities for those interested in pursuing the career of a data scientist.

If you are someone who is just starting out with data science, then you would like to know how to become a data scientist (<https://hackr.io/blog/how-to-become-a-data-scientist>) first.

## Data Science Interview Questions and Answers

However, if you're already past that and preparing for a data scientist job interview, here are the 50 top data science interview questions with answers to help you secure the spot:

**Question: Can you enumerate the various differences between Supervised and Unsupervised Learning?**

**Answer:** Supervised learning is a type of machine learning where a function is inferred from labeled training data. The training data contains a set of training examples.

Unsupervised learning, on the other hand, is a type of machine learning where inferences are drawn from datasets containing input data without labeled responses. Following are the various other differences between the two types of machine learning:

- **Algorithms Used** – Supervised learning makes use of Decision Trees, K-nearest Neighbor algorithm, Neural Networks, Regression, and Support Vector Machines. Unsupervised learning uses Anomaly Detection, Clustering, Latent Variable Models, and Neural Networks.
- **Enables** – Supervised learning enables classification and regression, whereas unsupervised learning enables classification, dimension reduction, and density estimation
- **Use** – While supervised learning is used for prediction, unsupervised learning finds use in analysis

Check here the detailed difference between Supervised Learning vs Unsupervised Learning (<https://hackr.io/blog/supervised-vs-unsupervised-learning>)

### **Question: What do you understand by the Selection Bias? What are its various types?**

**Answer:** Selection bias is typically associated with research that doesn't have a random selection of participants. It is a type of error that occurs when a researcher decides who is going to be studied. On some occasions, selection bias is also referred to as the selection effect.

In other words, selection bias is a distortion of statistical analysis that results from the sample collecting method. When selection bias is not taken into account, some conclusions made by a research study might not be accurate. Following are the various types of selection bias:

- **Sampling Bias** – A systematic error resulting due to a non-random sample of a populace causing certain members of the same to be less likely included than others that results in a biased sample.
- **Time Interval** – A trial might be ended at an extreme value, usually due to ethical reasons, but the extreme value is most likely to be reached by the variable with the most variance, even though all variables have a similar mean.
- **Data** – Results when specific data subsets are selected for supporting a conclusion or rejection of bad data arbitrarily.
- **Attrition** – Caused due to attrition, i.e. loss of participants, discounting trial subjects or tests that didn't run to completion.

### **Question: Please explain the goal of A/B Testing.**

**Answer:** A/B Testing is a statistical hypothesis testing meant for a randomized experiment with two variables, A and B. The goal of A/B Testing is to maximize the likelihood of an outcome of some interest by identifying any changes to a webpage.

A highly reliable method for finding out the best online marketing and promotional strategies for a business, A/B Testing can be employed for testing everything, ranging from sales emails to search ads and website copy.

**Question: How will you calculate the Sensitivity of machine learning models?**

**Answer:** In machine learning, Sensitivity is used for validating the accuracy of a classifier, such as Logistic, Random Forest, and SVM. It is also known as REC (recall) or TPR (true positive rate).

Sensitivity can be defined as the ratio of predicted true events and total events i.e.:

$$\text{Sensitivity} = \frac{\text{True Positives}}{\text{Positives in Actual Dependent Variable}}$$

Here, true events are those events that were true as predicted by a machine learning model. The best sensitivity is 1.0 and the worst sensitivity is 0.0.

**Question: Could you draw a comparison between overfitting and underfitting?**

**Answer:** In order to make reliable predictions on general untrained data in machine learning and statistics, it is required to fit a (machine learning) model to a set of training data. Overfitting and underfitting are two of the most common modeling errors that occur while doing so.

Following are the various differences between overfitting and underfitting:

- **Definition** - A statistical model suffering from overfitting describes some random error or noise in place of the underlying relationship. When underfitting occurs, a statistical model or machine learning algorithm fails in capturing the underlying trend of the data.
- **Occurrence** – When a statistical model or machine learning algorithm is excessively complex, it can result in overfitting. Example of a complex model is one having too many parameters when compared to the total number of observations. Underfitting occurs when trying to fit a linear model to non-linear data.
- **Poor Predictive Performance** – Although both overfitting and underfitting yield poor predictive performance, the way in which each one of them does so is different. While the overfitted model overreacts to minor fluctuations in the training data, the underfit

model under-reacts to even bigger fluctuations.

**Question: Between Python and R, which one would you pick for text analytics and why?**

**Answer:** For text analytics, Python (<https://hackr.io/tutorials/learn-python?ref=blog-post>) will gain an upper hand over R due to these reasons:

- The Pandas library in Python offers easy-to-use data structures as well as high-performance data analysis tools
- Python has a faster performance for all types of text analytics
- R (<https://hackr.io/tutorials/learn-r?ref=blog-post>) is a best-fit for machine learning than mere text analysis

**Question: Please explain the role of data cleaning in data analysis.**

**Answer:** Data cleaning can be a daunting task due to the fact that with the increase in the number of data sources, the time required for cleaning the data increases at an exponential rate.

This is due to the vast volume of data generated by additional sources. Also, data cleaning can solely take up to 80% of the total time required for carrying out a data analysis task.

Nevertheless, there are several reasons for using data cleaning in data analysis. Two of the most important ones are:

- **Cleaning data from different sources helps in transforming the data into a format that is easy to work with**
- Data cleaning increases the accuracy of a machine learning model

**Question: What do you mean by cluster sampling and systematic sampling?**

**Answer:** When studying the target population spread throughout a wide area becomes difficult and applying simple random sampling becomes ineffective, the technique of cluster sampling is used. A cluster sample is a probability sample, in which each of the sampling units is a collection or cluster of elements.

Following the technique of systematic sampling, elements are chosen from an ordered sampling frame. The list is advanced in a circular fashion. This is done in such a way so that once the end of the list is reached, the same is progressed from the start, or top, again.

**Question: Please explain Eigenvectors and Eigenvalues.**

**Answer:** Eigenvectors help in understanding linear transformations. They are calculated typically for a correlation or covariance matrix in data analysis.

In other words, eigenvectors are those directions along which some particular linear transformation acts by compressing, flipping, or stretching.

Eigenvalues can be understood either as the strengths of the transformation in the direction of the eigenvectors or the factors by which the compressions happens.

**Question: Can you compare the validation set with the test set?**

**Answer:** A validation set is part of the training set used for parameter selection as well as for avoiding overfitting of the machine learning model being developed. On the contrary, a test set is meant for evaluating or testing the performance of a trained machine learning model.

**Question: What do you understand by linear regression and logistic regression?**

**Answer:** Linear regression is a form of statistical technique in which the score of some variable Y is predicted on the basis of the score of a second variable X, referred to as the predictor variable. The Y variable is known as the criterion variable.

Also known as the logit model, logistic regression is a statistical technique for predicting the binary outcome from a linear combination of predictor variables.

**Question: Please explain Recommender Systems along with an application.**

**Answer:** Recommender Systems is a subclass of information filtering systems, meant for predicting the preferences or ratings awarded by a user to some product.

An application of a recommender system is the product recommendations section in Amazon. This section contains items based on the user's search history and past orders.

**Question: What are outlier values and how do you treat them?**

**Answer:** Outlier values, or simply outliers, are data points in statistics that don't belong to a certain population. An outlier value is an abnormal observation that is very much different from other values belonging to the set.

Identification of outlier values can be done by using univariate or some other graphical analysis method. Few outlier values can be assessed individually but assessing a large set of outlier values require the substitution of the same with either the 99th or the 1st percentile values.

There are two popular ways of treating outlier values:

1. To change the value so that it can be brought within a range
2. To simply remove the value

**Note:** - Not all extreme values are outlier values.

**Question: Please enumerate the various steps involved in an analytics project.**

**Answer:** Following are the numerous steps involved in an analytics project:

- **Understanding the business problem**
- Exploring the data and familiarizing with the same
- Preparing the data for modeling by means of detecting outlier values, transforming variables, treating missing values, et cetera
- Running the model and analyzing the result for making appropriate changes or modifications to the model (an iterative step that repeats until the best possible outcome is gained)
- Validating the model using a new dataset
- Implementing the model and tracking the result for analyzing the performance of the same

**Question: Could you explain how to define the number of clusters in a clustering algorithm?**

**Answer:** The primary objective of clustering is to group together similar identities in such a way that while entities within a group are similar to each other, the groups remain different from one another.

Generally, Within Sum of Squares (<https://discuss.analyticsvidhya.com/t/what-is-within-cluster-sum-of-squares-by-cluster-in-k-means/2706>) is used for explaining the homogeneity within a cluster. For defining the number of clusters in a clustering algorithm, WSS is plotted for a range pertaining to a number of clusters. The resultant graph is known as the Elbow Curve.

The Elbow Curve graph contains a point that represents the point post in which there aren't any decrements in the WSS. This is known as the bending point and represents K in K-Means.

Although the aforementioned is the widely-used approach, another important approach is the Hierarchical clustering. In this approach, dendrograms are created first and then distinct groups are identified from there.

### **Question: What do you understand by Deep Learning?**

**Answer:** Deep Learning is a paradigm of machine learning that displays a great degree of analogy with the functioning of the human brain. It is a neural network method based on convolutional neural networks (CNN).

Deep learning has a wide array of uses, ranging from social network filtering to medical image analysis and speech recognition. Although Deep Learning has been present for a long time, it's only recently that it has gained worldwide acclaim. This is mainly due to:

- **An increase in the amount of data generation via various sources**
- The growth in hardware resources required for running Deep Learning models

Caffe, Chainer, Keras, Microsoft Cognitive Toolkit, Pytorch, and TensorFlow are some of the most popular Deep Learning frameworks as of today.

### **Question: Please explain Gradient Descent.**

**Answer:** The degree of change in the output of a function relating to the changes made to the inputs is known as a gradient. It measures the change in all weights with respect to the change in error. A gradient can also be comprehended as the slope of a function.

Gradient Descent refers to escalating down to the bottom of a valley. Simply, consider this something as opposed to climbing up a hill. It is a minimization algorithm meant for minimizing a given activation function.

### **Question: How does Backpropagation work? Also, it state its various variants.**

**Answer:** Backpropagation refers to a training algorithm used for multilayer neural networks. Following the backpropagation algorithm, the error is moved from an end of the network to all weights inside the network. Doing so allows for efficient computation of the gradient.

Backpropagation works in the following way:

- **Forward propagation of training data**



- Output and target is used for computing derivatives
- Backpropagate for computing the derivative of the error with respect to the output activation
- Using previously calculated derivatives for output generation
- Updating the weights

Following are the various variants of Backpropagation:

- **Batch Gradient Descent** – The gradient is calculated for the complete dataset and update is performed on each iteration
- Mini-batch Gradient Descent – Mini-batch samples are used for calculating gradient and updating parameters (a variant of the Stochastic Gradient Descent approach)
- **Stochastic Gradient Descent** – Only a single training example is used to calculate gradient and updating parameters

### **Question: What do you know about Autoencoders?**

**Answer:** Autoencoders are simplistic learning networks used for transforming inputs into outputs with minimum possible error. It means that the outputs resulted are very close to the inputs.

A couple of layers are added between the input and the output with the size of each layer smaller than the size pertaining to the input layer. An autoencoder receives unlabeled input that is encoded for reconstructing the output.

### **Question: Please explain the concept of a Boltzmann Machine.**

**Answer:** A Boltzmann Machine features a simple learning algorithm that enables the same to discover fascinating features representing complex regularities present in the training data. It is basically used for optimizing the quantity and weight for some given problem.

The simple learning algorithm involved in a Boltzmann Machine is very slow in networks that have many layers of feature detectors.

### **Question: What are the skills required as a Data Scientist that could help in using Python for data analysis purposes?**

**Answer:** The skills required as a Data Scientist that could help in using Python for data analysis purposes are stated under:

1. Expertize in Pandas Dataframes, Scikit-learn, and N-dimensional NumPy Arrays.
2. Skills to apply element-wise vector and matrix operations on NumPy arrays.

3. Able to understand built-in data types, including tuples, sets, dictionaries, and various others.
4. It is equipped with Anaconda distribution and the Conda package manager.
5. Capability in writing efficient list comprehensions, small, clean functions, and avoid traditional for loops.
6. Knowledge of Python script and optimizing bottlenecks

**Question: What is the full form of GAN? Explain GAN?**

**Answer:** The full form of GAN is Generative Adversarial Network. Its task is to take inputs from the noise vector and send it forward to the Generator and then to Discriminator to identify and differentiate the unique and fake inputs.

**Question: What are the vital components of GAN?**

**Answer:** There are two vital components of GAN. These include the following:

1. **Generator:** The Generator act as a Forger, which creates fake copies.
2. **Discriminator:** The Discriminator act as a recognizer for fake and unique (real) copies.

**Question: What is the Computational Graph?**

**Answer:** A computational graph is a graphical presentation that is based on TensorFlow. It has a wide network of different kinds of nodes wherein each node represents a particular mathematical operation. The edges in these nodes are called tensors. This is the reason the computational graph is called a TensorFlow of inputs. The computational graph is characterized by data flows in the form of a graph; therefore, it is also called the DataFlow Graph.

**Question: What are tensors?**

**Answer:** Tensors are the mathematical objects that represent the collection of higher dimensions of data inputs in the form of alphabets, numerals, and rank fed as inputs to the neural network.

**Question: Why are Tensorflow considered a high priority in learning Data Science?**

**Answer:** Tensorflow is considered a high priority in learning Data Science because it provides support to using computer languages such as C++ and Python. This way, it makes various processes under data science to achieve faster compilation and completion within the stipulated time frame and faster than the conventional Keras and Torch libraries. Tensorflow supports the computing devices, including the CPU and GPU for faster inputs, editing, and analysis of the data.

**Question: What is Dropout in Data Science?**

**Answer:** Dropout is a toll in Data Science, which is used for dropping out the hidden and visible units of a network on a random basis. They prevent the overfitting of the data by dropping as much as 20% of the nodes so that the required space can be arranged for iterations needed to converge the network.

### **Question: What is Batch normalization in Data Science?**

**Answer:** Batch Normalization in Data Science is a technique through which attempts could be made to improve the performance and stability of the neural network. This can be done by normalizing the inputs in each layer so that the mean output activation remains 0 with the standard deviation at 1.

### **Question: What is the difference between Batch and Stochastic Gradient Descent?**

**Answer:** The difference between Batch and Stochastic Gradient Descent can be displayed as follows:

<b>Batch Gradient Descent</b>	<b>Stochastic Gradient Descent</b>
It helps in computing the gradient using the complete data set available.	It helps in computing the gradient using only the single sample.
It takes time to converge.	It takes less time to converge.
The volume is huge for analysis purpose	The volume is lesser for analysis purposes.
It updates the weight slowly.	It updates the weight more frequently.

### **Question: What are Auto-Encoders?**

**Answer:** Auto-Encoders are learning networks that are meant to change inputs into output with the lowest chance of getting an error. They intend to keep the output closer to the input. The process of Autoencoders is needed to be done through the development of layers between the input and output. However, efforts are made to keep the size of these layers smaller for faster processing.

### **Question: What are the various Machine Learning Libraries and their benefits?**

**Answer:** The various machine learning libraries and their benefits are as follows.

1. Numpy: It is used for scientific computation.

2. Statsmodels: It is used for time-series analysis.
3. Pandas: It is used for tubular data analysis.
4. Scikit learns: It is used for data modeling and pre-processing.
5. Tensorflow: It is used for the deep learning process.
6. Regular Expressions: It is used for text processing.
7. Pytorch: It is used for the deep learning process.
8. NLTK: It is used for text processing.

**Question: What is an Activation function?**

**Answer:** An Activation function helps in introducing the non-linearity in the neural network. This is done to help the learning process for complex functions. Without the activation function, the neural network will be unable to perform only the linear function and apply linear combinations. Activation function, therefore, offers complex functions and combinations by applying artificial neurons, which helps in delivering output based on the inputs.

**Question: What are the different types of Deep Learning Frameworks?**

**Answer:** The different types of Deep Learning Framework includes the following:

1. Caffe
2. Keras
3. TensorFlow
4. Pytorch
5. Chainer
6. Microsoft Cognitive Toolkit

**Question: What are vanishing gradients?**

**Answer:** The vanishing gradients is a condition when the slope is too small during the training process of RNN. The result of vanishing gradients is poor performance outcomes, low accuracy, and long term training processes.

**Question: What are exploding gradients?**

**Answer:** The exploding gradients are a condition when the errors grow at an exponential rate or high rate during the training of RNN. This error gradient accumulates and results in applying large updates to the neural network, causes an overflow, and results in NaN values.

**Question: What is the full form of LSTM? What is its function?**

**Answer:** LSTM stands for Long Short Term Memory. It is a recurrent neural network that is capable of learning long term dependencies and recalling information for the longer period as part of its default behavior.

**Question: What are the different steps in LSTM?**

**Answer:** The different steps in LSTM include the following.

- Step 1: The network helps in deciding the things that need to be remembered while others that need to be forgotten.
- Step 2: The selection is made for cell state values that can be updated.
- Step 3: The network decides as to what can be made as part of the current output.

**Question: What is Pooling on CNN?**

**Answer:** Pooling is a method that is used with the purpose to reduce the spatial dimensions of a CNN. It helps in performing downsampling operations for reducing dimensionality and creating pooled feature maps. Pooling in CNN helps in sliding the filter matrix over the input matrix.

**Question: What is RNN?**

**Answer:** The RNN stands for Recurrent Neural Networks. They are an artificial neural network that is a sequence of data, including stock markets, sequence of data including stock markets, time series, and various others. The main idea behind the RNN application is to understand the basics of the feedforward nets.

**Question: What are the different layers on CNN?**

**Answer:** There are four different layers on CNN. These include the following.

1. Convolutional Layer: In this layer, several small picture windows are created to go over the data.
2. ReLU Layer: This layer helps in bringing non-linearity to the network and converts the negative pixels to zero so that the output becomes a rectified feature map.
3. Pooling Layer: This layer reduces the dimensionality of the feature map.

4. Fully Connected Layer: This layer recognizes and classifies the objects in the image.

### **Question: What is an Epoch in Data Science?**

**Answer:** Epoch in Data Science represents one of the iterations over the entire dataset. It includes everything that is applied to the learning model.

### **Question: What is a Batch in Data Science?**

**Answer:** Batch is referred to as a different dataset that is divided into the form of different batches to help to pass the information into the system. It is developed in the situation when the developer cannot pass the entire dataset into the neural network at once.

### **Question: What is the iteration in Data Science? Give an example?**

**Answer:** Iteration in Data Science is applied by Epoch for analysis of data. The iteration is, therefore, classification of the data into different groups. For example, when there are 50,000 images, and the batch size is 100, then in such a case, the Epoch will run about 500 iterations.

### **Question: What is the cost function?**

**Answer:** Cost functions are a tool to evaluate how good the model performance has been made. It takes into consideration the errors and losses that are made in the output layer during the backpropagation process. In such a case, the errors are moved backward in the neural network, and various other training functions are applied.

### **Question: What are hyperparameters?**

**Answer:** Hyperparameter is a kind of parameter whose value is set before the learning process so that the network training requirements can be identified and the structure of the network can be improved. This process includes recognizing the hidden units, learning rate, epochs, and various others associated.

### **Question: Which skills are important to become a certified Data Scientist?**

**Answer:** The important skills to become a certified Data Scientist include the following:

1. Knowledge of built-in data types including lists, tuples, sets and related.
2. Expertize in N-dimensional NumPy Arrays.
3. Ability to apply Pandas Dataframes.
4. Strong hold over performance in element wise vectors.

5. Knowledge of matrix operations on NumPy arrays.

### **Question: What is an Artificial Neural Network in Data Science?**

**Answer:** Artificial Neural Network in Data Science is the specific set of algorithms that are inspired by the biological neural network meant to adapt the changes in the input so that the best output can be achieved. It helps in generating the best possible results without the need to redesign the output methods.

### **Question: What is Deep Learning in Data Science?**

**Answer:** Deep Learning in Data Science is a name given to machine learning, which requires a great level of analogy with the functioning of the human brain. This way, it is a paradigm of machine learning.

### **Question: Are there differences between Deep Learning and Machine Learning?**

**Answer:** Yes, there are differences between Deep Learning and Machine learning. These are stated as under:

<b>Deep Learning</b>	<b>Machine Learning</b>
It gives computers the ability to learn without being explicitly programmed.	It gives computers a limited to unlimited ability wherein nothing major can be done without getting programmed, and many things can be done without the prior programming. It includes supervised, unsupervised, and reinforcement machine learning processes.
It is a subcomponent of machine learning that is concerned with algorithms that are inspired by the structure and functions of the human brains called the Artificial Neural Networks.	It includes Deep Learning as one of its components.

### **Question: What is Ensemble learning?**

**Answer:** Ensemble learning is a process of combining the diverse set of learners that is the individual models with each other. It helps in improving the stability and predictive power of the model.

### **Question: What are the different kinds of Ensemble learning?**

**Answer:** The different kinds of Ensemble learning includes the following.

1. Bagging: It implements simple learners on one small population and takes mean for estimation purposes.
2. Boosting: It adjusts the weight of the observation and thereby classifies the population in different sets before the outcome prediction is made.

## Conclusion

That completes the list of the top data science interview questions. I hope you will find it useful to prepare well for your upcoming data science job interview(s).

Data Science is a top career profile nowadays. When looking for more Data Science Interview Questions, consider this popular udemy course: Data Science Career Guide - Interview Preparation (<https://click.linksynergy.com/deeplink?id=jU79Zysihs4&mid=39197&murl=https://www.udemy.com/course/data-science-career-guide-interview-preparation/>).

We also suggest you one of the best data science interview questions books: Practical Statistics for Data Scientists: 50 Essential Concepts 1st Edition (<https://geni.us/OTcA>).

Check out these best data science tutorials (<https://hackr.io/tutorials/learn-data-science?ref=blog-post>) to step up your data science game today.

Wish you good luck!

### People are also reading:

- Difference between Data Science vs Data Analytics (<https://hackr.io/blog/data-science-vs-data-analytics>)
- Get the Difference between Data Analyst vs Data Scientist (<https://hackr.io/blog/data-analyst-vs-data-scientist>)
- 10 Best Data Science Books (<https://hackr.io/blog/data-science-books>)
- Best Data Analytics Courses (<https://hackr.io/blog/data-analytics-courses>)
- Top 10 Python Data Science Libraries (<https://hackr.io/blog/top-data-science-python-libraries>)
- Top Data Science Interview Questions & Answers (<https://hackr.io/blog/data-science-interview-questions>)
- What is Data Analysis? (<https://hackr.io/blog/what-is-data-analysis>)
- What is Data Analytics? (<https://hackr.io/blog/what-is-data-analytics>)
- What is Data Analysis? (<https://hackr.io/blog/what-is-data-analysis>)









- Difference between Data Science vs Machine Learning (<https://hackr.io/blog/data-science-vs-machine-learning>)
- Top Bootstrap Interview Questions and Answers (<https://hackr.io/blog/bootstrap-interview-questions>)

 [Data science \(https://hackr.io/blog/tag/data-science\)](https://hackr.io/blog/tag/data-science) [Interview Questions \(https://hackr.io/blog/tag/interview-questions\)](https://hackr.io/blog/tag/interview-questions)

[Data Science Interview Questions \(https://hackr.io/blog/tag/data-science-interview-questions\)](https://hackr.io/blog/tag/data-science-interview-questions)

### Share:

  (<https://twitter.com/intent/tweet?text=Data+Science+Interview+Questions+https%3A%2F%2Fhackr.io%2Fblog%2Fdata-science-interview-questions>)  (<https://www.linkedin.com/shareArticle?mini=true&url=https://hackr.io/blog/data-science-interview-questions>)  (<http://www.reddit.com/submit?url=https://hackr.io/blog/data-science-interview-questions>)  (<https://news.ycombinator.com/submitlink?u=https://hackr.io/blog/data-science-interview-questions>)  (<https://api.whatsapp.com/send?text=https%3A%2F%2Fhackr.io%2Fblog%2Fdata-science-interview-questions>)



### Akhil Bhadwal

(<https://hackr.io/blog/author/akhilbhadwal>)

A Computer Science graduate interested in mixing up imagination and knowledge into enticing words. Been in the big bad world of content writing since 2014. In his free time, Akhil likes to play cards, do guitar jam, and write weird fiction. View all posts by the Author (<https://hackr.io/blog/author/akhilbhadwal>)

## Related Posts

(<https://hackr.io/blog/what-is-data-engineering>)

## **What is Data Engineering? How to Become a Data Engineer?**

(<https://hackr.io/blog/what-is-data-engineering>) **Read More** (<https://hackr.io/blog/what-is-data-engineering>)

(<https://hackr.io/blog/css-interview-questions>)

## **40+ Top CSS Interview Questions and Answers**

(<https://hackr.io/blog/css-interview-questions>) **Read More** (<https://hackr.io/blog/css-interview-questions>)

(<https://hackr.io/blog/google-interview-questions>)

### Google Interview Questions and Answers

(<https://hackr.io/blog/google-interview-questions>) **Read More** (<https://hackr.io/blog/google-interview-questions>)

## Leave a comment

**Email address\***

Your email will not be published

**Name\***

**Comment\***

**SUBMIT**

**Shayan Ali Bhatti**

Deep Learning is not just based on Convolutional Neural Network. It is based on neural networks, and if the hidden layers of neural networks are more than 3, we say that we are doing deep learning. Although some claim 3 or 5 hidden layers, but that is how deep learning is defined by pioneers of deep learning. Thanks for sharing the questions.

Reply

## Related Tutorials

 **OpenCV**

(<https://hackr.io/tutorials/learn-opencv>)

 **MATLAB**

(<https://hackr.io/tutorials/learn-matlab>)

 **OpenGL**

(<https://hackr.io/tutorials/learn-opengl>)

 **Data Science**

(<https://hackr.io/tutorials/learn-data-science>)

 **Artificial Intelligence**

(<https://hackr.io/tutorials/learn-artificial-intelligence-ai>)

## Recommended Learning

CS109 Data Science ([cs109.github.io](https://cs109.github.io))

(<https://hackr.io/tutorial/cs109-data-science>)

**Introduction to Data Science in Python** ([www.coursera.org](http://www.coursera.org))

(<https://hackr.io/tutorial/introduction-to-data-science-in-python>)

**IBM Data Science Professional Certificate** ([www.coursera.org](http://www.coursera.org))


(<https://hackr.io/tutorial/ibm-data-science-professional-certificate>)

**VIEW MORE** (<https://hackr.io/tutorials/learn-data-science>)

---

Blog (<https://hackr.io/blog>)   Roadmaps (<https://hackr.io/roadmaps>)   About Us (<https://hackr.io/about>)

Programming Tips (<https://chrome.google.com/webstore/detail/programming-tips/ooaiehbfgcjjeaiedpffeajkeikpl>)

Help & FAQ (<https://hackr.io/help>)   We  Feedback   Advertise (<https://hackr.io/advertise-with-us>)

(<https://play.google.com/store/apps/details?id=io.hackr.hackr&hl=en>)

(<https://apps.apple.com/in/app/hackr-io/id1188958684>)