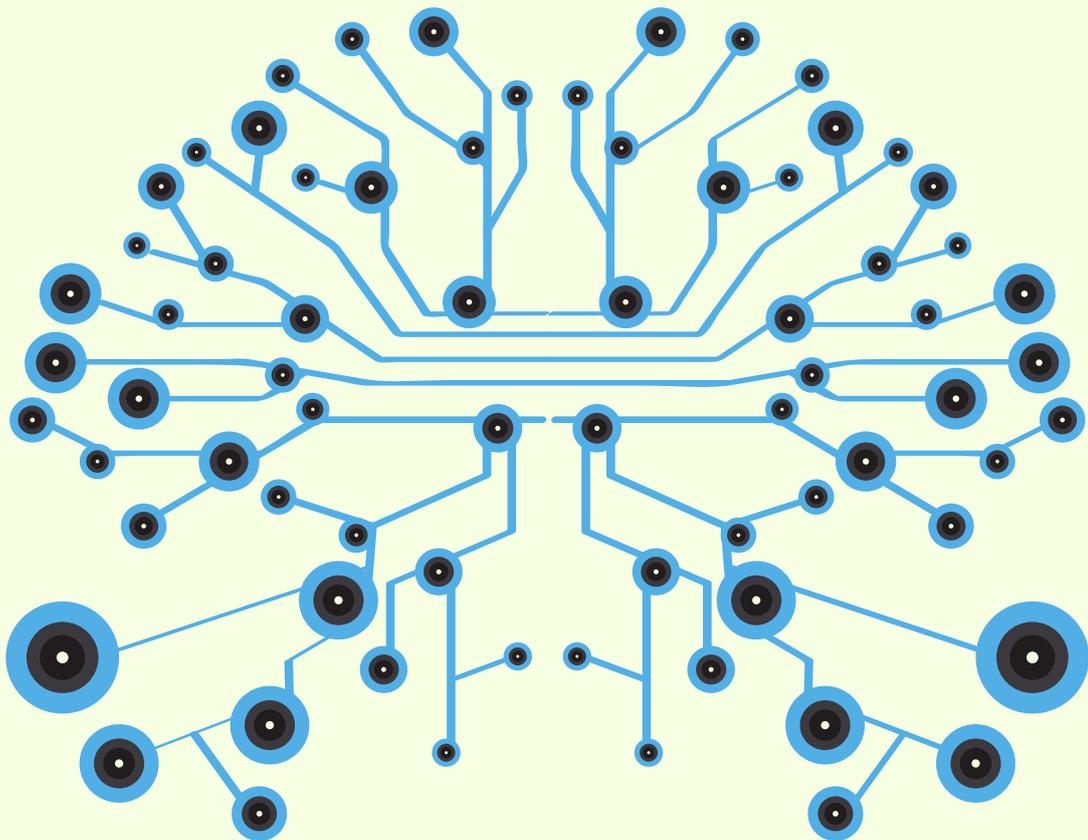


# DATA SCIENCE INTERVIEW PREPARATION GUIDE



NITIN MUKESH

*First Edition*

© 2021 Nitin Mukesh

All rights reserved. No portion of this book may be reproduced in any form without permission from the Author.

# TABLE OF CONTENTS

## **I. INTERVIEW PREPARATION**

- 1.1 How data science interview is different*
- 1.2 Data science campus recruitment Process*
- 1.3 Prepare Aptitude effectively*
- 1.4 Prepare for Data Science Interview*
- 1.5 Is that all asked?*
- 1.6 Some tips for interview preparation*

## **2. INTERVIEW QUESTIONS**

- 2.1 Statistics interview Question*
- 2.2 Regression analysis Questions*
- 2.3 Logistic regression Questions*
- 2.4 Machine learning Questions*
- 2.5 Coding Questions*
- 2.6 Guestiates Questions*
- 2.7 Puzzle Questions*
- 2.8 HR Questions.*

## **3. MISCELLANEOUS**

- 3.1 Tell me about yourself.*
- 3.2 How to be motivated during placement season?*

# How Data Science interview is different?



Clearing a data science interview isn't a big deal if you are well prepared. But the fact is that data science borrows concepts from many fields and that's why you need to have knowledge of all those fields to clear a data science interview.

Though there are different jobs related to data like data analyst, data engineer, Big data developer, all these domains are not the same.

Data scientist's job profiles differ from all these in a lot of ways.

There are a lot of concepts that you need to be aware of for clearing a data science interview.

Believe me, I have gone through a lot of interviews and failed many. The reason was that, if you lack in any of the areas, you will be discarded.

But, this doesn't make data science job interviews so complicated. What is required for a data science job, is already with you.

You just need to sharpen those skills and along with them, learn some new concepts.

This ebook is not for a detailed explanation of these topics, but it is about what to read, where to read from, and how to prepare for interview.

*Nitin Mukesh*

# Data Science Campus Recruitment Process



If I want to summarize the entire campus recruitment process for data science jobs, it can be divided into two sections:



- Technical Tests
- Interviews

In campus recruitment, or in off-campus jobs too, the first stage is to appear for a test. Those students who clear the test are asked to appear for an interview.

In the data science domain, the technical tests are mostly centered around common aptitude sections which include, quantitative aptitude, reasoning, sometimes English, or a few times coding.

Different companies have different types of tests. For these tests, you should prepare basic aptitude.

Coding tests are there in few companies only.

From a lot of tests that I have appeared in or seen, during my campus placement journey, here are few examples of which companies take what type of tests for their data science profile:

- **Aptitude:** Citi bank, ANZ, Wells Fargo, ICICI Lombard, IDFC First Bank, PharmaACE, Accenture, FCS Limited, General Mills,
- **Aptitude+Coding:** Delhivery, Goldman Sachs, Miko.ai, Flipkart, Icertis, Postman, TCS Research, Honda R&D,
- **Data Science-based test:** Microsoft, Micron Technology



Now, as you can see, most of the companies go for aptitude tests. Though the companies taking coding rounds visit some of the premier campuses only.

You should find out which company visits your college each year and what are their recruitment processes.

This is a really nice strategy to target a company that you want to be in.

You should always do your own research during the campus recruitment drive.

# How to prepare for Aptitude effectively?



To prepare for aptitude sections, you can use the following resources:

- **Quantitative aptitude by Dr. R. S. Agarwaal**
- **A Modern Approach To Verbal & Non-Verbal Reasoning By R.S. Aggarwal**
- **Quantitative Aptitude by Arun Sharma**
- **<https://www.indiabix.com/>**
- **<https://www.pariksha.co/placements>**

For beginners, Quantitative aptitude by R S Agarwal is really a good book. I have personally used this book for my aptitude preparations.

Most of the students feel comfortable with the aptitude section. Mostly aptitude can be solved without extra preparations.

But to get accuracy and speed, you can think of preparing for it too.

If you are thinking of preparing for aptitude from any book mentioned above, it can take a lot of time if you go solving each and every question.

I mean the Book By R. S. Agarwaal has more than 700 pages,

Reading the whole book will require a lot of time. Also, you need to prepare for the interview part. So, it is not possible to dedicate most of the preparation time to the aptitude section only.

But, there is a strategy that you can opt.

## 10-2-4 Rule



What I call it is the **10-2-4 Rule**.

So according to this rule, you have to follow the given steps each day:

- Read the formula given at beginning of a chapter, solve 10 questions from that chapter.
- From the previous two chapters solve 4 questions from each.

This way you will be able to go through all the chapters in 15-20 days. Now, each day solve few questions from 4-5 chapters at once.

Also, appear for the online mock test that you can get on the website given above.

Even if you feel confident, keep practicing few questions on regular basis.

What I have personally experienced is that in the beginning, we prepare aptitude very well, but as the campus placement goes on, if we are not selected in any company or any interview, we leave practicing aptitude and only focuses on interview.

Just solve a few questions. Maybe once a week you can do it so that in the next company you will be able to perform well again.

Enough about the aptitude, lets now move to the interview preparation section.

Now if you have cleared the aptitude round, here comes the interview. The interview questions are mainly centered on your resume, job profile, and your academic course(if you are from mathematics, statistics, computer science, or related fields).

But few topics are commonly asked if you are appearing for a data science interview. You should brush up on these basics concepts and also try to have some mock interviews in prior.

# How to prepare for a Data Science interview?



While preparing for data science interviews, you should take notes and also prepare from questions that have already been asked in companies in interviews.

For your reference, I have already listed the interview questions towards the end of this ebook. These questions have been given topic-wise along with the name of companies in which these questions have been asked.

Before mentioning the topics that are commonly asked in data science interviews, I would like to list some key points you must take care of while preparing for interviews.



- The first question that an interviewer is going to ask you will be, "**Tell me about yourself.**" Prepare this question really well. Later in this guide, I'll explain how to answer this question effectively.
- You should be well prepared with your resume. The interviewer doesn't know about you, what he has with you is your resume. So, be well prepared with your resume.

Now let's quickly reveal all those topics that you should be preparing for a data science interview.

If you are from Statistics, Computer Science, operation research, or data science-oriented degrees, chances are high that you might already be knowing most of the topics.

But, as I always say, don't forget to revise these. Most of the time we don't remember what we have studied a long back. So, without any hesitation, go back to revise all those concepts from your course notes.

I have also listed some of the books and courses from where you can read these topics.

Here are commonly asked topics:

## **Basic Statistic**

- Probability
- Bayes Theorem
- Mean, Median, Mode and its properties
- Random variables
- Skewness, Kurtosis
- Expectations and Variance with properties
- Correlation and Covariance

# Statistical Distributions

- Binomial distribution
- Poisson Distribution
- Geometric Distribution
- Exponential Distributions
- Normal Distributions
- Uniform Distributions
- Gamma and Beta Distributions

## Hypothesis Testing

- t-test
- f-test
- chi-square test
- ANOVA
- Interpretation of p-value
- Central Limit Theorem

Recommended book:

- **Fundamental of Mathematical statistics by Gupta and Kapoor**
- **An Introduction to Probability and statistics by Rohatgi and Saleh**

# Machine learning Algorithms

- Linear regression
- Logistic regression
- KNN
- SVM
- Decision Tree
- LDA and QDA
- Random Forest
- Bagging
- Boosting
- PCA
- Naive Bayes classifier
- Ridge Regression
- Lasso Regression
- Elastic Net
- K-Means clustering

Recommended book: **ISLR (Introduction to Statistical Learning)**

## Data Science Concepts

- Bias-variance tradeoff
- Precision and recall
- ROC-AUC curve
- Handling missing data
- Outlier handling
- Cross-validation
- Overfitting and Underfitting
- Regularization

Recommended book: **ISLR (Introduction to Statistical Learning)**

## SQL Concepts

- Joins in SQL
- Constraints in SQL
- Primary Key and Foreign Key
- Queries in SQL
- TRUNCATE, DELETE and DROP
- Use of Aggregate Functions
- WHERE, HAVING and GROUP BY clause

# Data Structure Concepts

- Array
- Linked List
- Stack
- Queue
- Binary search tree
- Binary Search
- Sorting algorithms

# Programming Concepts

- List, tuple, dictionary, and set
- Loop and function
- Recursion
- Dataframe
- Importing data
- Data Visualization Libraries
- ML Libraries

# Other concepts

- Puzzles
- Case study and Guestimates

# Is that all asked?



If you are appearing for a data science profile, most of the time, questions are based on the topics that have been mentioned above.

But, as I have told in the beginning that the data science profile nowadays has been diversified.

The company now hires for more specific roles in the data science domain, like Machine learning Engineer, Research Scientists, Time series analysts, banking companies like ANZ, Citi, Wells Fargo, IDFC First bank, etc hires for Modelling risk analysts or financial analysts.

Other product-based companies hire NLP engineers, Computer vision engineers, Image processing specialists, etc, which are deep learning specified roles.

But, deep learning is an entirely different domain. This ebook is only for data science specialized roles.

In data science specified roles, the interview will be around these topics+your resume+ Job profile.

# Tips for preparing a data science interview



- **Make a nice resume:** A resume is very important for any job interview. Resume-making can be a tedious task, but a good resume is not made in a single day.

You will have to sit for many days, making small changes, consulting with your senior or some experts, and ask for advice.

Make sure that your resume doesn't look overqualified or underqualified.

You may be wondering, why an overqualified resume is bad?

Well, being “overqualified” is seen as a bad thing, not because someone “overqualified” will perform too well, but because they are expected to be dissatisfied with the job.

Granted, a person who is “overqualified” will generally learn their job in a minimal time.

If you are in an overqualified zone, you might be over-qualified by the interviewer and can have a bad interview.

I have been rejected from a lot of interviews due to being over questioned.

*Nitin Mukesh*

So, don't make this mistake. Spend some time making a good resume. Take help from others. It takes some time to make an awesome eye-catching resume.

- **Improve your communication skills:** You may be knowing a lot of things in data science, have done awesome projects, and are well-versed with everything, but you are not able to explain yourself, guess what? You will be rejected.

It is a harsh truth that the company prefers one with good communication skills rather than one with complete knowledge.

And if you have both, then you are gold.

***But, what if you are not able to explain yourself in the interview?***

This can be fixed. You need to learn to explain things better.

Record yourself over the phone and watch it, listen to your recorded voice, see how it sounds. Have questions discussion with your friends. Discuss in the group. Have a group mock interview with your friends in which one will explain the answers to a group.

With practice, you can easily master this skill. But, don't ignore it. What students commonly think is, I am an expert in data science, nothing will happen if my communication skills are bad.

But, guess what happens?

The interviewer will give preference to those who can explain. So, start to prepare early, work on your weak points, and do outstanding in any interview.



- **Engage yourself in the interview:** Don't just sit answering what the interviewer asks you. The most important thing is you have to be present in the interview.

When told, ask questions from the interviewer. Ask how their journey has been till now. When you don't know the answer, ask if they can give some additional hints and then try to solve. This will certainly have a positive impact on the interviewer.

- **Don't interrupt the interviewer:** If the interviewer asks anything and you know the answer very well, then don't interrupt in between. Let the interviewer complete the question.

After the interviewer finishes, you can start. I had this bad habit of interrupting. But, I came to know about this mistake during one of the mock interviews.

Also, if you don't know the answer to any of the questions, then you can politely say that "Sir, I don't know the answer." or "I have not come across this term till now."

You are not supposed to know everything. I myself was only able to answer 85% of questions during the interview for the company I was placed in.

My one of the batchmate who was placed at highest package from our department was able to answer 80% of the questions.

The key is that, whatever you explain, be clear, be specific, use examples and tell an amazing story out of these.

- **Start simple, be clear:** When asked any questions, explain in simple and clear language. Don't complicate things.

Don't explain things in hurriedness, explain slowly. If you need time to think, tell the interviewer, "Sir, can I take one minute to think?"

Simplicity is key.

- **Be stress-free:** If you will be stress-free during the entire interview, you will be able to think clearly and answer better.

I still remember my first face-to-face interview for Accenture. It was for a data science internship. Though I was very much worried and anxious. But, as the interview proceeds, I become more comfortable.

But, due to COVID-19, the entire placement procedure was online. There was a problem with me. I could not think focus and recall what I had learned in an online interview. I usually used to get distracted and forget things.

I had found this thing before the placement procedure was about to begin.

I used to practice a lot with my laptop camera on. Also, I used to have online mock interviews with my friends. These things helped me a lot, and I started to feel comfortable.

But as placement proceeded, failing a lot of interviews really taught me how to handle pressure during interviews. I

It's all part of learning.

- **Analyze yourself:** During preparations, you should analyze yourself at each step, like what you need to study, find your weak spots, work on them.

I advise maintaining a record of what you have to read, what is still left, and with time update it. Keeping a record of what you studied really makes things easier.

After the interview also, write what was asked, what were the questions that you weren't able to answer, search on the internet, and read them.

This way, even if you are not selected, now you have some experience. That's how ML models also learn. With more data, the models get better.

**Do your own research:** If you are selected for an interview, read about the company. Do your own research. Find about what the company does, how are their work environment, how is the working culture.

You will find some questions about the company that HR may ask in HR Round.

Also, if you appear for any start-ups, do find about them. Sometimes some fake companies do visits and later revoke your job offer.

At IIT Bombay, a company that was offering 75 lakhs CTC with 20 LPA gross package on day 14. Everyone was shocked that any company is offering this much on the second last day.

But, soon my friends informed me that this company seems to be fake. We did much online research and found that this company is really fake and had nothing to offer.

We immediately informed the Placement cell, and then the company was questioned. It was ultimately black-listed.

You know the funny part is the company has only one employee who was HR, Founder, Interviewer, and everything.

I know it may sound shocking that such companies dare to visit IIT Bombay for Placement. But, in the COVID situation, the fake companies do try to take advantage. So, be aware of that.

Always do your own research.

- **Help others and take help:** When you get placed, help your friends in placement. The best part of life is when you start helping each other.

You should really be grateful for your part and help others.

You will really find yourself happy when your friends will also be placed.

Also, don't hesitate to take help from your friends if you need it. I was stupid that I took it so late.. But believe me, put your ego aside and take help. Give mock interviews to them.

Life is really awesome. Ego and jealousy don't have a place in our life—enjoy the placement journey.



# Data Science Interview Questions



# STATISTICS INTERVIEW QUESTIONS

**ASKED IN DATA  
SCIENCE  
INTERVIEWS**



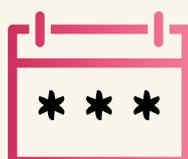
# Commonly Asked Topics

- **Probability Theory**
- **Sampling Theory**
- **Hypothesis Testing**
- **Distribution Theory**
- **Time series Analysis**
- **Design of Experiment**

- Prove  $E(x+y)=E(x) + E(y)$  (**Delhivery**)
- What strong and weak stationarity? (**ICICI Securities**)
- Explain ARIMA forecasting step by step. (**ICICI Securities**)
- What is CLT, z-test, t-test, F test? (**ICICI Securities**)
- When is a t-test used instead of a z-test? (**ICICI Securities**)
- What do you mean by Replication and Randomization? (**IPSOS MMA**)
- Trick question on the application of chi-square test. (**ICICI Securities**)
- Test of equality of proportions. (**ICICI Securities**)
- Structure of a box plot. (**ICICI Lombard**)
- What does correlation mean? What type of relationship can be found by this? (**ANZ Bank**)
- What is the importance of stationarity in a time series? (**ANZ Bank**)
- What is skewness in data and how will you remove it? (**LafargeHolcim**)
- Role of R squared, r, F test in determining features in Machine learning.
- What is stationary time series and why it is important? (**Cognizant**)
- Define Randomised Block design? Where do we use RBD? (**Cognizant**)

- What is stratified random sampling and how it is different from systematic random sampling? **(Cognizant)**
- What is the p-value? Explain with the help of an example related to marketing research. **(IPSOS MMA)**
- What are the assumptions of ANOVA? **(JSS, ANZ)**
- explain CRD (Completely Randomized Design)? **(IPSOS MMA)**
- What is AB testing? **(ICICI Securities)**
- Consider two fair dice that have been rolled. What is the probability of obtaining sum 8?**(IPSOS MMA)**
- The lengths of a professor's classes have a continuous uniform distribution between 50.0 min and 52.0 min. If one such class is randomly selected, find the probability that the class length is more than 51.7 min. **(64squares)**
- Difference between stratified and cluster sampling. **(JSS)**
- When we use censoring and what are type-1 and type-2 censoring in Reliability. **(JSS)**
- ANOVA is generally used when we want to test the significance of the difference between more than two independent sample means. Why we don't use pairwise t-test in this case? **(JSS, Accenture)**

- What are type-1 error and p-value in inference? **(JSS, Accenture)**
- What do you mean by hypothesis testing? **(Eli Lilly)**
- There are 3 ants sitting on three corners of a triangle. All ants randomly pick a direction and start moving along the edge of the triangle. What is the probability that any two ants collide? **(Accenture, Delhivery)**
- You have 9 red balls and 1 black ball in a bag. You are picking balls 10 times with replacement. What is the probability of getting at least one black ball from the bag? **(Accenture)**
- Explain Central Limit Theorem. **(ANZ, Wipro)**
- What is the normal distribution and what is the shape of the curve? **(WNS Global)**
- What is the relationship between media, mode and mean in positively skewed data? **(ANZ)**
- What are statistic and data science? **(ANZ)**
- You have to estimate the average height of the population of India, how will you go about it? **(ICICI Lombard)**
- What is the use of binomial distribution and its PMF? **(Wipro)**
- What do you mean by selection bias? **(Wipro)**



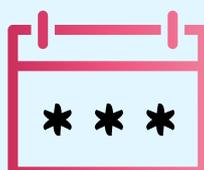
# REGRESSION ANALYSIS INTERVIEW QUESTIONS

ASKED IN REAL  
INTERVIEWS



1. What are the assumptions of linear regression?  
**(Delhivery, ANZ bank, Citi Bank, Accenture)**
2. What is the meaning of multicollinearity? **(ANZ, Amazon)**
3. How to detect multicollinearity? **(Amazon, Delhivery)**
4. What do you understand by VIF (Variance Inflation Factor)? **(Amazon)**
5. What is the difference between R-squared and adjusted R-squared? **(Delhivery, ANZ bank, Citi Bank, Accenture)**
6. How to deal with multicollinearity in data? **(Citi, Accenture)**
7. Explain forward and backward elimination?  
**(Accenture)**
8. How PCA works? **(ICICI securities, Amazon, Miko.ai)**
9. Explain Ridge and Lasso Regression? **(Delhivery, ANZ bank, Citi Bank, Accenture, Amazon)**
10. Can SVM be used for regression? **(Miko.ai)**
11. What is the curse of dimensionality? Can you give an example?
12. What is the difference between the coefficient of determination, and coefficient of correlation?
13. Give methods of variable selection in Regression Analysis? **(Delhivery, ANZ bank, ICICI securities)**

14. Why do we perform the residual analysis? **(ANZ)**
15. What are L1 and L2 penalization? **(Miko.ai)**
16. What is heteroscedasticity? How does it affect the regression coefficients? **(ANZ)**
17. Why does only  $VIF > 10$  implies that there is multicollinearity, why not choose  $vif > 8$ ? **(IDFC First Bank)**
18. In my dataset, if I have 100 observations and 1500 features, do you think whether I would be able to fit the regression model onto that or not? **(IDFC First Bank)**
19. For a single variable, how will you detect outliers? **(ICICI Lombard)**
20. How correlation between two variables will change in presence of an outlier? Will it increase, decrease or remain constant? Explain how, using its formula. **(ICICI Lombard)**
21. What are influential and leverage points? Which of them have more effect on the model? **(ANZ, Wells Fargo)**
22. Does multicollinearity impact the prediction of a machine learning algorithm? **(Wells Fargo)**



# LOGISTIC REGRESSION INTERVIEW QUESTIONS

ASKED IN  
PLACEMENT



# Logistic Regression

- Logistic regression is a supervised Machine learning algorithm used for classification.
- It is mainly used to model a binary dependent variable.



- Logistic Regression can be easily used to classify more than 2 classes.
- Example of Binary Classification,
  - To predict whether an email is a spam or not
  - Whether the tumor is malignant or not

- What is the difference between linear and logistic regression? **(ANZ. Wells Fargo, Accenture)**
- Can you use Logistic regression for regression problems? **(Accenture)**
- Explain the working of Logistic Regression? **(ANZ, Delhivery)**
- What do you mean by Generalized Linear Model? **(ICICI Lombard)**
- Is logistic regression a Generalized Linear Model?
- How Support Vector Machine different from Logistic regression?
- Write the expression for the Logistic function.
- What do you mean by deviance in Logistic regression? **(ANZ)**
- What are penalized logistic regression models? Why did you use them? How are they better than logistic regression models? **(FCS Limited)**
- Differentiate between Lasso, Ridge, and Elastic Net. **(Wells Fargo)**
- How to handle imbalanced data in classification? Which metric to use in this case? Why is accuracy not a good measure in this case? **(FCS Limited)**
- How will you deal with overfitting in the case of Logistic Regression? **(Delhivery)**
- Interpretation of parameters of the logistic regression model. **(ANZ)**

- Steps to build a Logistic Regression model from scratch. **(ANZ)**
- If the accuracy of the model is 95%, is it good? **(ANZ)**
- Difference between Logistic regression and decision tree. **(Delhivery)**
- Difference between Logistic regression and Random Forest. **(ANZ)**
- What do you mean by cost function? **(Amazon)**
- What is a convex function? **(Amazon)**
- Explain Gradient Descent Algorithm. **(Meru cab)**
- How will you use logistic regression if your data has more than 2 classes?



# INTERVIEW QUESTIONS ON CLASSIFICATION ALGORITHMS

ASKED IN DATA  
SCIENCE INTERVIEWS



# Commonly Asked Topics

- **Logistic regression**
- **KNN**
- **SVM**
- **Decision Tree**
- **Random Forest**
- **Bagging**
- **LDA & QDA**
- **Naive Bayes classifier**
- **Classification Metrics**
- **Overfitting and underfitting**

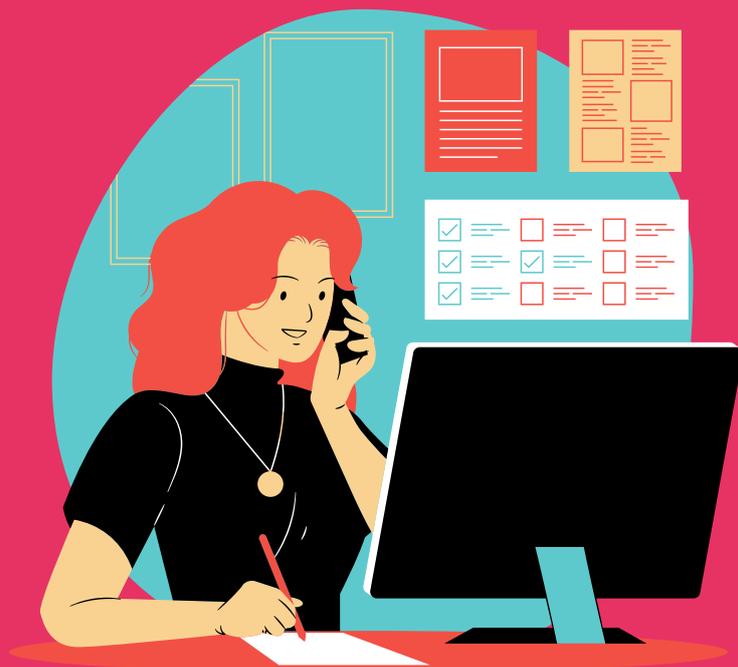
- Difference between logistic and SVM. **(Accenture)**
- How to deal with overfitting in Logistic Regression? **(Delhivery)**
- If the accuracy of the model is 95%, is it good? **(ANZ)**
- Does standardization have any impact on the performance of the model in the case of a random forest? **(Wells Fargo)**
- How to handle imbalanced data in classification? Which metric to use in this case? Why is accuracy not a good measure in this case? **(FCS Limited, Citi)**
- Explain the Logistic regression model. **(ANZ, Miko.ai, IPSOS)**
- Suppose you have 200 variables, will you consider all the variables to build a model? **(ANZ)**
- Difference between random forest and logistic regression? **(ANZ)**
- What is a random forest and how it differs from Adaboost? **(Meru cabs)**
- Explain bagging in Machine learning? **(TCS innovation lab)**
- Explain F score, precision, recall, ROC curve. **(Meru cabs)**
- When can you say, KNN model is overfitted? **(Delhivery)**

- Write the loss function of logistic and linear regression. **(Delhivery)**
- How is the prediction made in the random forest? **(FCS Limited)**
- How is XGBoost different from gradient boosting? **(Wells Fargo)**
- What will happen if we use MSE in logistic regression? **(Delhivery)**
- If you have categorical features in your data, then how will you find the Nearest Neighbor using KNN? **(ANZ)**
- Explain the working of the Support Vector Machine Algorithm? **(ICICI Lombard, Cognizant)**
- What is the kernel trick in SVM? Why is it computationally efficient? **(FCS Limited)**
- Explain the complete working of the decision tree algorithm. **(WNS Global, TCS innovation lab)**
- What is a confusion matrix? **(ANZ)**
- Difference between Linear regression vs logistic regression. **(Wells Fargo)**
- What is underfitting and overfitting for example? **(Accenture)**
- What are the steps to build a model from scratch? **(ANZ)**
- Why random forests are called random? **(FCS Limited)**

- Can Ridge or Lasso be used for NLP? (**Miko.ai, NLP**)
- What is the difference between Bagging and Boosting? (**Miko.ai, TCS Innovation Lab**)
- Differentiate between SVM and Random Forest? (**Miko.ai**)
- Which is better, Random forest or decision tree? (**TCS Innovation Lab**)
- What is the Gini index and entropy in the decision tree? (**Cognizant**)
- What do you mean by deviance in Logistic regression? (**ANZ**)
- What is a convex function? (**Amazon**)
- What are the criteria to evaluate a Logistic regression model? (**MMA IPSOS**)



# CODING QUESTIONS ASKED IN DATA SCIENCE INTERVIEWS



# Commonly Asked Topics

- **Loops**
- **Functions**
- **Recursion**
- **Dynamic Programming**
- **Sorting algorithm**
- **Binary Search**

Coding questions are commonly asked in data science interviews. But, if you are sitting for banking companies like ANZ, Citi, ICICI, IDFC first bank, etc, you can expect no or minimal coding questions.

Companies like Delhivery, Accenture, TCS, Flipkart, Microsoft, etc ask a little bit more in-depth questions from coding sections.

But, if you are comfortable with the topics, I have mentioned above, then you are good to go. More is better, but at least prepare that many of things.

I bet you can easily get these topics from YouTube. If you are interested in reading these topics in-depth, the channel that I recommend is given below:

- Codebasics: **[Link](#)**

Don't worry if you are not an expert in coding. As I told you, something is better than nothing. At least read the basics if you can't master all these topics.

Many times, questions based on the application of the basics of loops are asked. So, at least prepare the basics.

Here are some questions that have been asked in the data science and machine learning engineer interview.

- Write a code to find unique integers and the count of an array. **(Ignitarium Solutions)**
- Explain unsigned and signed integer representation in binary. **(Ignitarium Solutions)**
- You have a list of natural numbers. Using that, make a list of even numbers in just 1 line of code using python. **(FCS Limited)**
- How to make a function with variable number of parameters in python? **(FCS Limited)**
- How do you merge data frames in Pandas? **(FCS Limited)**
- What is the use of groupby function in Python? **(FCS Limited)**
- You have a list of names of cities. A city can occur more than once on the list. Write a Python code to count the number of each city. **(Accenture)**
- You are given two sorted lists you have to merge them in a single list in sorted order in  $O(n)$ . **(Accenture)**
- There are  $n$  steps in a stairs. You can hop either one or two steps at a time. Write a code to find the number of ways one can reach the  $n$ th step, given  $n$ . **(Accenture)**
- What is pandas used for? **(TCS Innovation lab)**
- What is a data frame? **(TCS Innovation lab)**

- How will you create an empty data frame in Python? **(TCS Innovation labs)**
- What are the different types of loops in programming? **(TCS Innovation labs)**
- How will you detect an infinite loop in a while loop? **(TCS Innovation labs)**
- Write code for Binary Search. **(Delhivery, Cargill)**
- How do we delete an element from a linked list? **(TCS Innovation labs)**
- Write a code to find the mean and standard deviation of a given list of numbers, without using libraries. **(AIRA Matrix)**
- How did you import a module from the library in python? **(LafargeHolCim)**
- What is the difference between an array and a linked list? **(TCS Innovation labs)**
- Write a recursive and non-recursive code for preorder tree traversal. **(Cargill Data Engineer)**



# GUESTIMATES ASKED IN REAL INTERVIEWS

ASKED IN  
PLACEMENT



# Why Guesstimates?

- We'll start by discussing guesstimates, for which candidates are asked to come up with a figure, usually the size of a market or the number of objects in an area or to estimate a particular data.



- Although guesstimates are more commonly given in interviews for consulting positions, they do pop up in finance and analytics interviews as well.
- Practicing guesstimates is a good way to begin preparing for stress questions in finance interviews, as they force candidates to think aloud – precisely what interviewers want to see.

# How to solve?

- The most important thing to remember about brainteasers, guesstimates, or even simple math questions that are designed to be stressful is to let your interviewer see how your mind works. And, that's what your interviewer wants to see.
- The best approach for a guesstimate question is to think of a funnel. You begin by thinking broadly, then slowly narrowing down the situation towards the answer. Show how you reach to answer.



- For any guesstimate or brainteaser question, you will need to understand whether your interviewer will be providing any direction or whether you will have to make assumptions.

# An Illustration

- e.g. Let us say you have to estimate the number of trees on your college campus.



- You can start by asking few initial questions to the interviewer like, will I have to estimate full-grown trees or small trees too?
- This shows you are trying to understand the situation.
- Then proceed with dividing your campus area into two parts, heavily populated with trees and less populated with trees.
- Then, you can make an initial estimate that let us say per 100 square meters in densely populated areas, 40 trees are there and hence calculate total trees by dividing the area of the densely populated region. The same goes for the other one.

# Conclusion

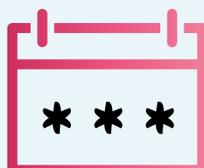
- See, you don't have to be exact in your answer. But, you have to show how can you break down a problem to solve it efficiently. That is what interviewer is interested in.



- With this, let's see some of the guestimate questions asked in real finance and analytics interviewers by different companies.

- Amount of revenue from the mobile industry in India **(KPMG)**
- Estimate the number of trees on the IIT Bombay campus.
- Number of Maggi consumed on campus in a month. **(Deloitte)**
- Number of smokers in Hyderabad **(Deloitte)**
- Number of Vada Paos served in the XYZ Mess **(Deloitte)**
- Number of Maggi packets sold on the first day of the relaunch **(Deloitte)**
- Number of flights taking off in a day from the Delhi Airport **(Deloitte)**
- The number of taxis in Bombay. **(Deloitte)**
- Number of people who have ever lived on the Earth **(Deloitte)**
- Number of vehicles in a toll gate of your choice **(Amazon)**
- Number of calls to our customer care in a week **(Amazon)**
- Number of people using FB on campus **(Amazon)**
- The market for cricket bats in the country **(Amazon)**
- The market for leather shoes **(Amazon)**
- Estimate the number of people on Patna Gandhi Maidan. **(ANZ)**
- Market size for electric insect repellent **(KPMG)**
- Estimate market size of electric bulbs in India **(KPMG)**
- Guesstimate number of daily flights in India **(KPMG)**
- Estimate market size of online retail in India **(Amazon)**
- Estimate the number of office chairs in Delhi **(EY)**
- Estimate number of autos in India **(Accenture)**
- Estimate number of Gmail app users in India **(Accenture)**

- Number of ticket counters required for Lucknow metro station **(Amazon)**
- Titan is launching a watch priced at Rs. 12,000. Estimate market size **(Accenture)**
- Royal Enfield is launching a new bike priced at 2 lakh in India. Estimate market size. **(Deloitte)**
- Estimate Revenue of Make my Trip in a year **(Deloitte)**
- Number of people at Lucknow airport in a day **(i3)**
- Number of cabs required at Lucknow airport **(i3)**
- Calculate the total amount of revenue from the mobile industry in India **(KPMG)**
- Estimate the amount of food wasted inside IIT Bombay hostels. **(Accenture)**
- You have to estimate the average height of the population of India, how will you go about it? **(ICICI Lombard)**
- Suppose the vaccine for Corona is ready. How will you plan its delivery? **(ICICI Lombard)**
- Estimate the number of masks which are disposed of every day in a city. **(ANZ Bank)**
- Estimate the number of flights going from Mumbai airport in a week? **(ANZ Bank)**
- Estimate the number of persons you need to run the campaign at the airport? **(ANZ Bank)**



# PUZZLES FOR DATA SCIENCE INTERVIEW

ASKED IN  
PLACEMENT



# How to answer puzzle questions?

When the interviewer asks you a puzzle, you don't need to tell an answer directly or come to conclusion at once.

Explain how you arrive at the answer. Explain the step-by-step procedure of getting to the answers. There are a lot of puzzles that are commonly asked.

I'll list some sources from where you can read puzzles.

Try to read 2-3 puzzles daily. Don't see the answers directly, just try to think for some time and then see the answer and analyze if you are correct.

I used to read 3-4 puzzles daily before going to sleep. This way I was able to cover most of the puzzles in 2 months.

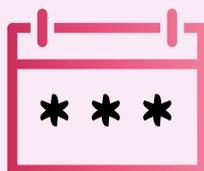
Here are some sources from where you prepare puzzles:

- Logically Yours (YouTube Channel): [Link](#)
- GeeksForGeeks Puzzles: [Link](#)

Here are few puzzles that have been asked in data science interviews:

- The lengths of a professor's classes have a continuous uniform distribution between 50.0 min and 52.0 min. If one such class is randomly selected, find the probability that the class length is more than 51.7 min. **(64 squares)**
- There are 3 ants sitting on three corners of a triangle. All ants randomly pick a direction and start moving along the edge of the triangle. What is the probability that any two ants collide? **(Accenture, Delhivery)**
- You have 9 red balls and 1 black ball in a bag. You are picking balls 10 times with replacement. What is the probability of getting at least one black ball from the bag? **(Accenture)**
- Suppose that we wish to know which stories in a 100-story building are safe to drop eggs from, and which will cause the eggs to break on landing. What strategy should be used to drop eggs such that the total number of drops in the worst case is minimized and we find the required floor? **(Delhivery)**

- You are provided with twenty-five different horses and you must find out who are the fastest horses. You can conduct a race of five horses only at one time. There is no point in the race where you can find out the actual speed of a horse in a race. How many races will it take to help you determine the fastest three horses? **(ANZ)**
- There is an ant on the top corner of a room and it has to travel to the diagonally opposite bottom of the room? What's the shortest way possible? **(ANZ)**
- Given two hourglass of 4 minutes and 7 minutes, the task is to measure 9 minutes. **(ICICI Lombard)**



# HR QUESTIONS FOR DATA SCIENCE INTERVIEW

ASKED IN  
PLACEMENT



# Why HR Questions?

After the initial rounds of candidates selection, every company holds an HR interview round to get to know the potential candidates personally.



An HR interview's main aim is to estimate the candidates' personality, background, strengths, and weaknesses and assess whether or not one is a good fit for a specific role.

In other words, the HR round decides if you'll be called in for the following rounds.

Though there can be a variety of questions that can be asked in the HR round depending on the job role and your degree. But, here are some questions that are more common in any interview:

- Tell me about yourself.
- Why should we hire you?
- How do you see yourself in the next 5 years?
- What are your strengths and weaknesses?
- Why do you want to join this company?

These 5 questions are the most common interview questions asked in the HR round. But that's not all.

To give you a glimpse of what kind of questions can be asked in the HR round, I have listed some questions below for your reference. These questions have been asked in different companies like Accenture, EY, ANZ, TCS, Citi bank, Wells Fargo, ICICI, UHG, etc.

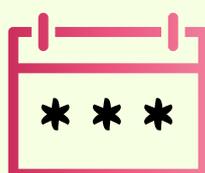
You don't need to write down answers to all these questions. Just discuss some of these questions with your friends during preparation.

The key is to be well prepared.

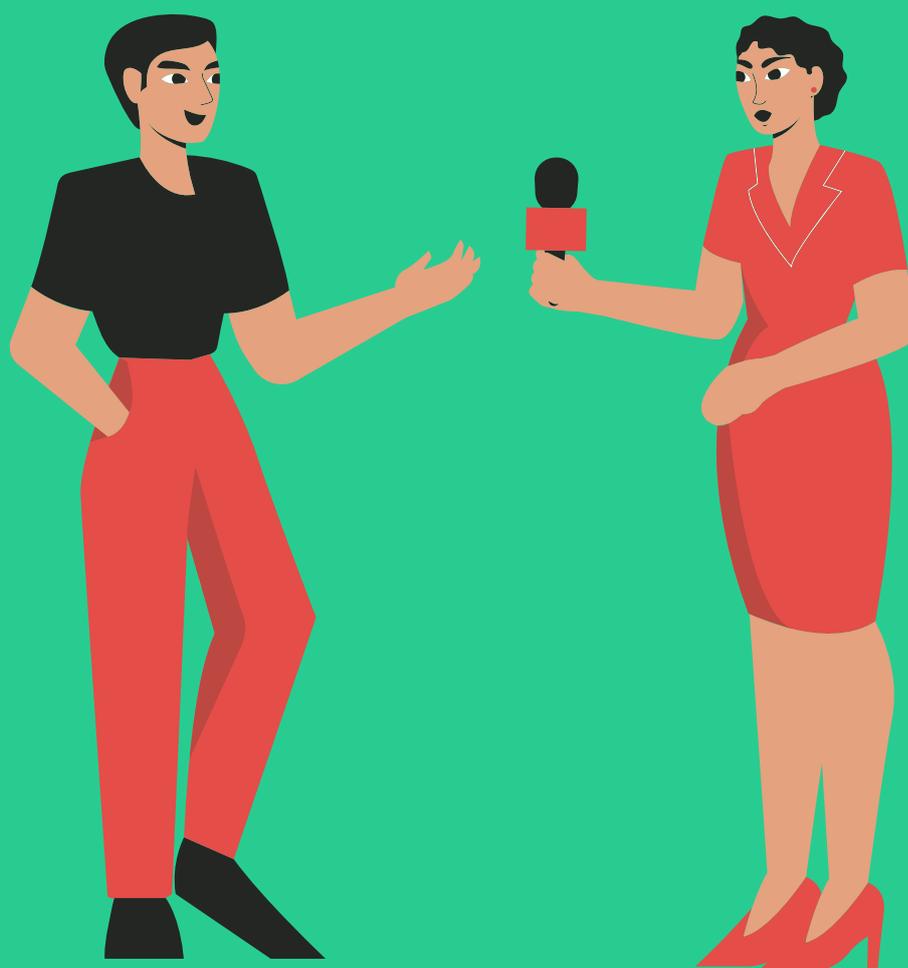
- Tell me something about yourself.
- How's your technical interview, rate yourself out of 10?
- Why statistics/data science?
- Have you ever used data science in your real life or any day-to-day application of data science?
- What is the biggest problem in India, how can you contribute to it as a data analyst?
- What are your expectations from our company?
- What are your long-term plans?
- What are your strengths and weakness? (at least 3)
- How are you doing to overcome your weaknesses?
- What are your hobbies?
- Where do you want to see yourself after 5 years in our company?
- According to you, a big reputed company matter or high salary in a start-up?
- What do you prefer - Teamwork or individual work
- What are the ethics, an employee should possess?
- Do you have managerial skills?
- Do you any issues regarding job location?
- What is your greatest fear?
- Do you have any questions from us?

- You are from statistics, why do you want to join IT Sector?
- What is your main motivation behind this profile?
- Any failure in life and what did you learn from it and how you succeeded afterward?
- What do you understand from this job profile?
- What's your aim in life?
- Where do u see urself in 10 years not from a carrier perspective but as a personality?
- Something u tried very hard to achieve but couldn't achieve?
- 2 things that you like about your college?
- 2 reasons that make you suitable for this profile?
- Why do you want to join a bank?
- Would u be ready to move abroad if given the opportunity
- Why are you interested in joining the finance sector?
- How will you deal with the problem of no background in finance?
- What makes you wake up in the morning?
- Suppose your co-worker working on a project with you drops the project in between. How would you make sure that the project is completed before its deadline?

- Why choose the corporate sector after having a good academic report?
- What are your thoughts on corporate life?
- How do your friends define you?
- If you have infinite time and money what you'll do?
- Have you come across a situation when you have a proposal and your council members are not ready to listen to you?
- How do you manage events?
- What steps you have taken when everything in your life is not going well?
- Suppose one company is offering you the work you like to do and with a decent salary and another company gives you a higher salary but you don't like the work in that company then which company you will choose?
- Tell me about how you feel when you think in the past that if you would have improved certain kinds of things then the result could have been better and give an example.



# Tell me about yourself.



It is the most important question in the interview. This is the first question you will be asked. Though, if asked in the technical round, the answer will be different than what you say if asked in HR Round.

I still remember my first interview for an internship. When the interviewer asked me, tell me about myself, I was not able to answer. But, somehow I managed to tell about my educational background and all. Then I was blank.

Later I came to know that all the students prepared for this question prior to the interview. I was stupid there who had no idea about it. So, prepare this question really well.

Here are the main components of the answer:

- Educational Background
- Projects and internships
- Your strength
- Conclusion

Let us see how can you make an ideal answer to this question.

So, here is a sample answer to this question:

Educational  
Background

"Thanks for giving me this opportunity to introduce myself. I have done my graduation from BHU in computer science and my MSc from IIT Bombay in Applied Statistics. **(You may tell about your school education too)**

Projects and  
Internships

If I talk about my professional experience and projects, I have done a machine learning research internship at IIIT Allahabad on self-taught learning using autoencoder.

I have done three projects, My MSc project was on Statistical analysis of proteins of COVID-19 patients, My other project was on the prediction of cancer mortality rate for US countries based on the application of Regression analysis.

My third project is on-time series and statistical analysis of the covid-19 scenario in Pune. **(You can talk about competitions you have won or Position of Responsibility here)**

Your  
strength

I'm proficient in Data Science, Machine learning, and coding concepts using Python. I am good at picking new technologies fast and I think time management is my strength.

Conclusion

And I feel I am now ready for a more challenging assignment and the prospect of working with XYZ firm really excites me a lot."

*Well, this is the answer that I told during my interview at TCS.  
"Awesome introduction, Nitin." the Interviewer replied.*

Prepare your answer well, record it and listen to it again and again so that it should sound perfect. Make sure that it should be between 60 seconds to 90 seconds.

The answer to the same questions will be different for the HR Round.

Here are the main components of this answer:

- Educational Background
- Strength
- Hobbies
- Interests
- Conclusion

Education background and conclusion will be the same. In the strength part, you have to tell about your strength like time management, fast learners, dedication, etc, then your hobbies and interests like music, sports, traveling, etc.

Yes, the ideal answers should lie between 60-75 seconds.

Prepare two types of answer for this question in advance and nail the interview.

**There is no  
failure except in  
no longer trying.**



**How to be motivated during  
placement season?**

You prepare for the campus placement, but not everyone gets selected in the very first company.

But, you don't need to get disheartened 😊 . The campus interview sometimes can be full of ups and downs. You don't know which is your lucky day.

Despite being prepared, I was not able to clear the interview during my campus placement season. I was able to clear most of the coding and aptitude tests, but in the interview round, I was getting rejected.

Sometimes the interviewer praised me for the interview round, but the company didn't respond.



During those hard times, it feels like everything is such a waste. What you have been doing till now is of no use.

You see your peers getting placed and here still you are waiting for another chance. Maybe someone with no experience or less technical skills than you is getting placed and still, you are not.

Self-doubt starts to kick in and you will start thinking that maybe you are not made for this job. All those lucky people got selected and you didn't.

But believe me, there is a lucky day for you. You can go through all sadness. It's the consequences of what is happening to us.

You need to keep going. Prepare even more. Analyze your mistakes. Improve them. Sometimes, there may not be your fault, accept it.

The road to success is not easy and not fair for everyone. After each interview, write a short note on how you feel about this interview? Think if you need to prepare more. Write those questions that you couldn't answer. Revise those topics that you were not able to explain clearly.

Improve yourself. You have to potential to get through it. Everyone has to go through this pain sometimes.

I had signed 110 companies, selected for tests in 40+ companies, interviewed 17 times before I got placed as an ML Researcher at TCS. And there is still someone who got placed even later.

I remember a story my brother was telling me when I was continuously getting rejected from interviews.

"There was a student in his college. His coding and aptitude skills were good. He had cleared tests and appeared for 100+ companies during college placement before he got placed. "

Keep going. You are going to learn a lot when one day you will be placed.

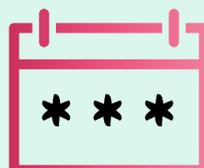
I'm able to write these things because I have experienced the mindset of a student who is getting rejected in every interview.

The key to an interview is simple: Prepare well, present yourself even better.

Put yourself 100% in the interview. Be confident and explain clearly.

That's all.

All the best for your next data science job interview!!!



**For more Data science placement preparation resources, Please scan:**



**I would like to ask you for a moment to fill this feedback form:**

**<https://forms.gle/JNMLr6zK6yYAo7UG9>**