1. **How does Big Data Analytics help increase the revenue of a company?**

Data Analytics helps the companies of today’s world in numerous ways. Following are the foundational concepts in which it helps:

• Effective use of data to relate to structured growth

• Effective customer value increase and retention analysis

• Manpower forecasting and improved staffing methods

• Bringing down the production cost majorly

1. **In your opinion, what does a Data Engineer majorly do?**

A Data Engineer is responsible for a wide array of things. Following are some of the important ones:

• Handling data inflow and processing pipelines

• Maintaining data staging areas

• Responsible for ETL data transformation activities

• Performing data cleaning and the removal of redundancies

• Creating ad-hoc query building operations and native data extraction methods

1. **What are some of the technologies and skills that a Data Engineer should possess?**

Following are the important technologies that a Data Engineer must be proficient in:

• Big data technologies such as Apache Spark, Data Lake, Data Factory

• Python

• SQL

Followed by this, a Data Engineer must also have good problem-solving skills and analytical thinking ability.

1. **What is the difference between a Data Architect and a Data Engineer?**

A Data Architect is a person who is responsible for managing the data that comes into the organization from a variety of sources. Data handling skills such as database technologies are a must-have skill of a Data Architect. The Data Architect is also concerned with how changes in the data will lead to major conflicts in the organization model.

Now, a Data Engineer is the person who is primarily responsible for helping the Data Architect with setting up and establishing the Data Warehousing pipeline and the architecture of enterprise data hubs.

1. **Have you earned any sort of certification to boost your opportunities as a Data Engineer?**

Yes, I have earned the certification ”Big Data Cloud Engineer” on completion of ”The Ultimate Big Data Master’s Program (Cloud Focused)” course which covered

Big data fundamental concepts

Datalake fundamentals

Distributed processing

High level APIs Dataframe and Spark SQL

Pyspark optimization

Azure Databricks, Data Factory and Synapse

I am currently preparing to get the certification in Azure as well

1. **Do you have any experience working in the same industry as ours before?**

This question is a frequent one. It is asked to understand if you have had any previous exposure to the environment and work in the same. Make sure to elaborate the experience you have, with the tools you’ve used and the techniques you’ve implemented. This ensures to provide a complete picture to the interviewer

1. **Why are you applying for the Data Engineer role in our company?**

Here, the interviewer is trying to see how well you can convince them regarding your proficiency in the subject, handling all the concepts needed to bring in large amounts of data, work with it, and help build a pipeline. It is always an added advantage to know the job description in detail, along with the compensation and the details of the company, thereby, obtaining a complete understanding of what tools, software packages, and technologies are required to work in the role.

1. **What is your plan after joining this Data Engineer role?**

While answering this question, make sure to keep your explanation concise on how you would bring about a plan that works with the company set up and how you would implement the plan, ensuring that it works by first understanding the data infrastructure setup of the company, and you would also talk about how it can be made better or further improvised in the coming days with further iterations.

1. **Do you have prior experience working with Data Modeling?**

If you are interviewed for an intermediate-level role, this is a question that you will always be asked. Begin your answer with a simple yes or no. It is alright if you have not worked with data modeling before, but make sure to explain whatever you know about data modeling in a concise and structured manner. It would be advantageous if you have used tools like Pentaho or Informatica for this purpose.

1. **Tell me about yourself.**

**What they’re really asking: What makes you a good fit for this job?**

Ofcourse! My name is Vanathi and I am currently working as a java developer at GE solutions. I have had a fulfilling career as a Java developer, but my true passion lies in data engineering.

What really sparked my interest in the field of data engineering was the project I worked on during my previous role in NHS. We were dealing with a significant amount of patient’s data and I was responsible for managing and organising it to derive insights for decision making. I became passionate about the idea of designing and optimizing data pipelines, ensuring data quality, and making data accessible and usable for various stakeholders. I started to explore data engineering concepts and began taking online course “The Ultimate Big Data Masters Program (Cloud Focused) 8 months course” to further develop my skills. The more I learned, the more I was drawn to the technical challenges and the opportunity to contribute to data-driven solutions. Additionally, I was inspired by industry trends and success stories where data engineering played a pivotal role in transforming businesses. It became clear to me that data engineering is my career choice and I'm eager to pursue it wholeheartedly.

I have developed strong skills in python, sql, pyspark, azure based big data technologies such datafactory, databricks, ADLS gen2 which I believe will be valuable in the world of data engineering. I am committed to continually learning and growing in this field and I am confident that my unique perspective and dedication to excellence will make me a valuable addition to your data team.

I am particularly excited about the opportunity to work as a data engineer at your company because of your company’s reputation for data driven decision making and innovative data solutions.

1. **Why did you choose to pursue a career in data engineering?**
2. **Describe your path to becoming a data engineer?**
3. **What is a data engineer’s role within a team or company?**

What they’re really asking: What is a data engineer responsible for?

For this question, recruiters want to know that you’re aware of the duties of a data engineer. What do they do? What role do they play within a team? You should be able to describe the typical responsibilities, as well as who a data engineer works with on a team. If you have experience as a data scientist or analyst, you may want to describe how you’ve worked with data engineers in the past.

The interviewer might also ask:

• What do data engineers do?

• How do data engineers work within a team?

• What impact does a data engineer have?

1. When did you face a challenge in dealing with unstructured data and how did you solve it?

What they’re really asking: How do you deal with problems? What are your strengths and weaknesses?

Essentially, a data engineer’s main responsibility is to build the systems that collect, manage, and convert raw data into usable information for data scientists and business analysts to interpret. This question aims to ask about any obstacles you may have faced when dealing with a problem, and how you solved it.

This is your time to shine, where you can describe how you make data more accessible through coding and algorithms. Rather than explaining the technicalities at this point, remember the specific responsibilities listed in the job description and see if you can incorporate them into your answer.

The interviewer might also ask:

• How do you solve a business problem?

• What is your process for dealing with and solving problems during a project?

• Can you describe a time when you encountered a problem and solved it in an innovative manner?

Data engineer process questions

Most often, data engineer job candidates will be asked about their projects. If you’ve never been a data engineer previously, you can describe projects that you either worked on for a class or one you posted on GitHub, a code hosting platform that promotes collaboration among developers.

1. Walk me through a project you worked on from start to finish.

What they’re really asking: How do you think through the process of acquiring, cleaning, and presenting data?

You’ll definitely be asked a question about your thought process and methodology for completing a project. Hiring managers want to know how you transformed the unstructured data into a complete product. You’ll want to practice explaining your logic for choosing certain algorithms in an easy-to-understand manner, to demonstrate you really know what you’re talking about. Afterward, you’ll be asked follow-up questions based on this project.

The interviewer might also ask:

• What was the most challenging project you’ve worked on, and how did you complete it?

• What is your process when you start a new project?

1. What algorithm(s) did you use on the project?

What they’re really asking: Why did you choose this algorithm, and can you compare it with other similar algorithms?

They want to know how you think through choosing one algorithm over another. It might be easiest to focus on a project that you worked on and link any follow-up questions to that project. If you have an example of a project and algorithm that relates to the company’s work, then choose that one to impress the interviewer. List the models you worked with, and then explain the analysis, results, and impact.

The interviewer might also ask:

• What is the scalability of this algorithm?

• If you were to do the project again, what would you do differently?

1. What tools did you use on the project?

What they’re really asking: How did you arrive at your decision to use certain tools?

Data engineers must manage huge swaths of data, so they need to use the right tools and technologies to gather and prepare it all. If you have experience using different tools such as Hadoop, MongoDB, and Kafka, you’ll want to explain which one you used for that particular project.

You can go into detail about the ETL (extract, transform, and load) systems you used to move data from databases into a data warehouse, such as Stitch, Alooma, Xplenty, and Talend. Some tools work better for back-end, so if you can communicate strong decision-making abilities, then you’ll shine as a candidate who’s confident in their skills.

The interviewer might also ask:

• What are your favorite tools to use, and why?

• Compare and contrast two or three tools that you used on a recent project