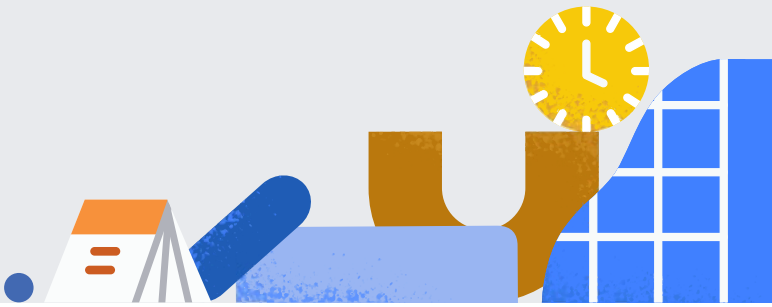
FACEBOOK

DATA ENGINEERING

Onsite Interview Guide



What You'll Find in This Guide

Welcome to your prep guide for your data engineer onsite interview at the Facebook company. Our data engineers put together this guide so you know what to expect and how to prepare.

Interview overview

What will the format and timing of your onsite interview look like?

Prepare to be on the Facebook campus for approximately five hours during your onsite interview. You'll be meeting with five or more different people, and the interview will be structured into the following (in no particular order):

- **Three Technical Interviews:** 1 hour each
- **Ownership Interview:** 30 minutes
- **Lunch:** 1 hour

Prep for your 3 technical interviews

How to prepare for your technical interviews

The three technical interviews collectively will assess your skills for product sense, analysis, data modeling SQL/ETL and coding.. You'll be using a whiteboard to communicate ideas, write actual SQL and code, draw schemas, and visualize data and process flow. If you haven't done a whiteboard interview before, you should consider practicing at home with an actual whiteboard, or pen and paper.

Product sense / analysis prep

At Facebook, we expect data engineers to have not only a strong technical aptitude but also a keen product sense. All three technical interviews will be case studies of typical product challenges that we solve with data. Your interviewer will assess your ability to think critically about the needs of the product in each scenario and how you translate those needs into a robust technical solution.

One way you can practice this is to think about metrics that companies similar to Facebook might use. For example, you can look at their financial statements to see what metrics they deem strategically important. From there, you should think through how you would calculate those metrics, and what you would do if they started moving unexpectedly.

- **Facebook Quarterly Results**
- **Facebook Products**

Data modeling prep

You'll brainstorm the data needs of a user product. Then you'll design a data mart to support analytics use cases and write select SQL statements to produce specific results.

To practice data modeling, go through the main products of several large tech companies, such as Facebook, LinkedIn or Amazon. Ask yourself, how would you model each function you use from the organization's product? Create logging designs for how you think data should be captured, then design data models to support analytical queries and reporting needs for those products. Here are some resources which might help you with prep work for data modeling:

- **Data Modeling Tutorial**
- **Data Modeling in Big Data**
- **Data Model**
- **Data Mart**
- **Dimensional Modeling**
- **Denormalization**

SQL prep

We'll ask you to focus on basic SQL constructs. Practice all types of joins, aggregate functions, analytical functions, set operators, and subqueries. You can take the work you did in data modeling, and practice how you'd load and transform the data from the logging sources into your target tables using SQL. During this exercise, also think through how you might help improve efficiency and scalability for processing large data volumes. Use the links below to practice:

- **SQL Tutorial**
- **PostgreSQL Exercises**
- **Netflix Case Study**
- **Insight on Data Swarm, Facebook's Custom ETL Tool**

Coding prep

Make sure that you know all the data structures and how to manipulate them well. Familiarize yourself with string, sets operations, etc., in your programming language of choice, and make sure you understand how dictionaries, lists, and loops work. Here are some resources to help you prepare, depending on which language you plan to use during the interview:

- **Python Guide**
- **Java Guide**
- **Scala Guide**

Ownership prep

What is your ownership interview?

Ownership

Data engineers need to take initiative and influence their fellow data engineers and XFN partners in their role, so we'll ask questions about your past experiences where you've been able to demonstrate this. The videos below will help give you a better sense of what drives the data engineers within Facebook Analytics.

- **An Inside Look at Data Engineering at Facebook**
- **Fireside Chat with a Data Engineering Director**
- **Data Engineering at Facebook: Bringing Communities Together**

5 Tips for your onsite interview

Things to keep in mind during your interviews

1. Listen for hints.

Interviewers might ask questions such as, "Are you sure you want to use a loop there?" This gives you an opportunity to consider alternatives to how to solve the problem and to demonstrate that you can learn quickly on the spot and implement feedback.

2. Think out loud.

It helps your interviewer follow along and learn about your problem-solving skills. They'll want to understand why you're making certain decisions. (For example: Why subqueries instead of joins?)

3. Clarifying questions.

Make sure you're asking clarifying questions as you go along (there won't be tricks but you'll need to ensure you have all the information you need). For instance, be comfortable asking, "Would you like me to keep going?"

4. Speed and efficiency.

Speed and efficiency are important. Think more about breadth over depth because it's important for the interviewer to get a signal on all your skillsets, not just one deep signal on one subject.

5. Share why you want to work at Facebook.

Your interviewer will want to know specific reasons rooted in your experience and what you're passionate about.

Appendix / resources

Links to exercises, information, and guides to help you prepare

Our team collected some helpful resources with content and activities for your onsite interview. Take a look through the list as you prepare.

Product Sense

- **How to Deep Dive in Your Product Funnel Performance with GA and Data Studio**
- **Most Important Product Metrics**
- **How Would You Measure the Success or Failure of a Product Feature?**

Data modeling

- **Data Modeling Tutorial**
- **Data Modeling in Big Data**
- **Data Model**
- **Data Mart**
- **Dimensional Modeling**
- **Denormalization**

SQL

- **SQL Tutorial**
- **PostgreSQL Exercises**

Coding

These resources contain basic procedural concepts for Python, Java, or Scala. Use the guide for the language you plan to use during the interview.

- **Python Guide**
- **Java Guide**
- **Scala Guide**

Data engineering at Facebook

At Facebook, data engineering plays a critical role in our product development strategy. Facebook is incredibly data-driven, and every feature, launch decision, and product opportunity is tested and proven with data.

Data engineers play a critical role in this effort by embedding within cross-functional teams where they partner closely with other functions like product managers, software engineers, designers, researchers, data scientists, etc. Working together, these cross-functional teams are collectively responsible for building products and services to serve billions of people around the world.

As a function, data engineering operates within the analytics organization alongside data science. This strong partnership enables us to build data assets which unlock the in-depth analysis we need to drive our products forward. Generally speaking, data engineers provide all the data; data scientists use that data to explain why things are happening. Analytics eventually rolls up into general management of the product.

Data engineers at Facebook need to have keen product awareness and the ability to hold a strategic conversation about the product, its possibilities, and its future. They need to be capable of anticipating future needs and designing data systems today to reduce the time and complexity involved in meeting those needs tomorrow. Successful data engineering produces better decisions, and subsequently, better outcomes for Facebook.

By nature, data engineers are going to receive many requests. As a result, prioritization is a critical skill requirement. Data engineers work in a partnership model and not as a service. As such, they need to strike the right balance between supporting the day-to-day needs of the product team and working on big bets. Facebook's culture prioritizes autonomy; independent decision-making is an essential factor in achieving success.

Technically speaking, the day-to-day job of a data engineer includes a wide variety of focus areas. Data engineers are responsible for high-quality logging, either through direct implementation or through a strong instrumentation partnership with software engineering. They are also responsible for complex integration work, bridging data from various sources together, enriching it, and operationalizing it via core tables, aggregate tables for convenient access, and dashboards for consumption. Data engineers also make data sets available for experimentation systems (A/B testing frameworks), training machine learning models, and sometimes delivering data directly back into Facebook's products.

On the Data Engineering team, there are no project managers, business analysts, ETL engineers, or visualization developers. Instead, there are solely data engineers who take on those roles and work end-to-end across the stack. But that isn't to say we don't specialize. We have a variety of archetypes that allow for deep specialization in a given area. Facebook is a strengths-based company, and we always aim to align each person's unique strengths with the right team and projects.

- Below are some videos and bios of key members from the Data Engineering team. These will help give you a better sense of what drives the data engineers within Facebook Analytics.
    - **An Inside Look at Data Engineering at Facebook**
    - **Byte Size Bio: Rohit G.**
    - **Byte Size Bio: Michael B.**
    - **Byte Size Bio: Shawna J.**
    - **Chat with Data Engineering Director Vish A.**
    - **How to Get Users and Grow**

Thanks for taking the time to review this guide and good luck in the interview - you'll do great!