Kamran Ali
@aTechGuide

How to break into Data Engineering in 2023

Let's look into Actionable Tips

atech.guide

Idea is NOT to learn
All the tools

But to build a base toolkit

If we demonstrate good, workable knowledge

Its a good indicator that we will Learn other tools / tech with the Job

Learn the tools in given
Sequence

1. Python and Git

1. Python is a interpreted, Object Oriented Programming Language

2. Git is a distributed version control System

3. We will use Python to write and schedule batch Data Pipelines.

4. As a Proof of work, publish code on Github so that we have something tangible to show in resume.

2. MySQL and SQL

1. We need a Data Store to Read data from and to Write data to. Workable understanding of a Database is really helpful

2. SQL is used to Query the Database

3. Learn the basic DDL, DML Operations in MySQL / Postgres.

4. Later you may add one NoSQL database to the mix e.g. Cassandra or MongoDB

3. Spark

1. Spark is a Distributed Compute Technology which provide APIs to transform data on a single or group of machines.

2. We can use Spark to write Scalable / fault tolerant Data Pipelines.

3. While Learning Spark, you're writing code in Python, using SQL for transformation (as spark SQL) and Pushing code in GitHub. So those Skills are still getting Practiced.

atech.guide

4. Airflow

1. Airflow is used to Develop, Schedule and Monitor Batch Oriented workflows

2. We will use Airflow to schedule Data Pipelines written on Spark

3. In Airflow we write DAGs, which are writing in Python.

4. Continue pushing code to GitHub

After 6 - 7 months, we are good
to start appearing for interviews

What about cloud e.g. AWS / Azure?

I will still say, we're good to appear
for Interviews

While giving interviews identify the gaps in what you know and what's being asked and work on it.

Optionally, you may start with AWS but it should not pause your interviews

All the BEST

Let me know in Comments, if you have any questions
or you disagree

Kamran Ali
@aTechGuide

That's a Wrap

1. Follow Me on LinkedIn for more valuable content - I'm Kamran Ali

2. Turn on the bell Notification in my profile