## Data Engineering Road Map(Links Are **Underlined)**

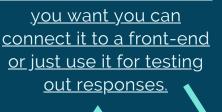


THINGS YOU SHOULD BE LEARNING ALONG THE WAY

CI/CD, Logging and Monitoring Tools, Debugging Tools(and the general skill set)

#### 2.BUILD YOUR FIRST PROJECT

Build A basic Flask API. If <u>you want you can</u> or just use it for testing out responses.



1.CODING BASICS

SQL

Python

be bad to learn some

server concepts like

SFTP. Firewalls. etc)

3. Linux(Also it wouldn't

### SECOND PROJECT

4. BUILD YOUR

- 1. Scrape an online source 2. Store encrypted data into SFTP
- 3. Create dimensional model
- 4. Pull data from SFTP and load into Data Warehouse(Don't worry too much about Workflows just yet)



#### 5.START LEARNING **ABOUT TESTING**

Unit testing, Integration Testing, etc Check out TDD Course Like <u>Unit TDD For Python</u>

#### 3.LEARN ABOUT DATA WAREHOUSES

- Read Kimball's Data Warehouse Guide
- 2. Take The Basics To Data Warehousing Udemy
- 3. Read An Introduction: Is Data Engineering For You

#### **5.LEARNING ABOUT** SOME WORKFLOW TOOLS

Launch Airflow with Docker. This will allow you to get a high-level understanding of Docker while learning Airflow.

#### **6.LEARN ABOUT CLOUD** AND NOSQL

- 1. Cloud Data Warehouses and ETLs using Googles Data Certificate or Azure's
- 2. <u>Check Out FreeCodeCamps -</u> Database Systems - Cornell University Course (SQL, NoSQL, Large-Scale Data Analysis)

#### 8.START STUDYING FOR INTERVIEWS

Start practicing some DS&A as well as SQL and data warehousing <u>questions</u>

# 10.1 FARN **ENOUGH**

Pick 1 dashboarding tool and learn how it function. Also learn a little about dashboarding UI/UX.

UI/UX AND

DASHBOARDING



#### 7. LEARN ABOUT STREAMING AND DISTRIBUTED SYSTEMS

- 1. Consider checking out Andrea's videos Why Learn Kafka
- 2. Take Frank Kane's Course on Spark and Hadoop
- Check Out StartDataEngineerings Projects



#### 9.BUILD YOUR THIRD PROJECT

<u>Use a Cloud Managed</u> Service like AWS Kinesis to stream data and pair it with a batch ETL tool that pulls from a second source.

You can store all the data in BigQuery and create some form of Data viz. This will also he;p you learn the next step



#### 11.PICK SOME OF YOUR OWN

At this point, you probably have found some personal favorites in terms of tool types. Why not learn a little more about them?

If you liked Docker, learn about Kubernetes, if you liked Dashboarding, pick a different tool, if you like the cloud, then why not learn a new cloud stack?

Some links may be affiliates that are helping support my channel. Thank you!



Check out Felipe Hoffa For Some Great Examples Of Documentation

