

Member-only story

# Fundamentals of Data Engineering Book — Key Learning Points

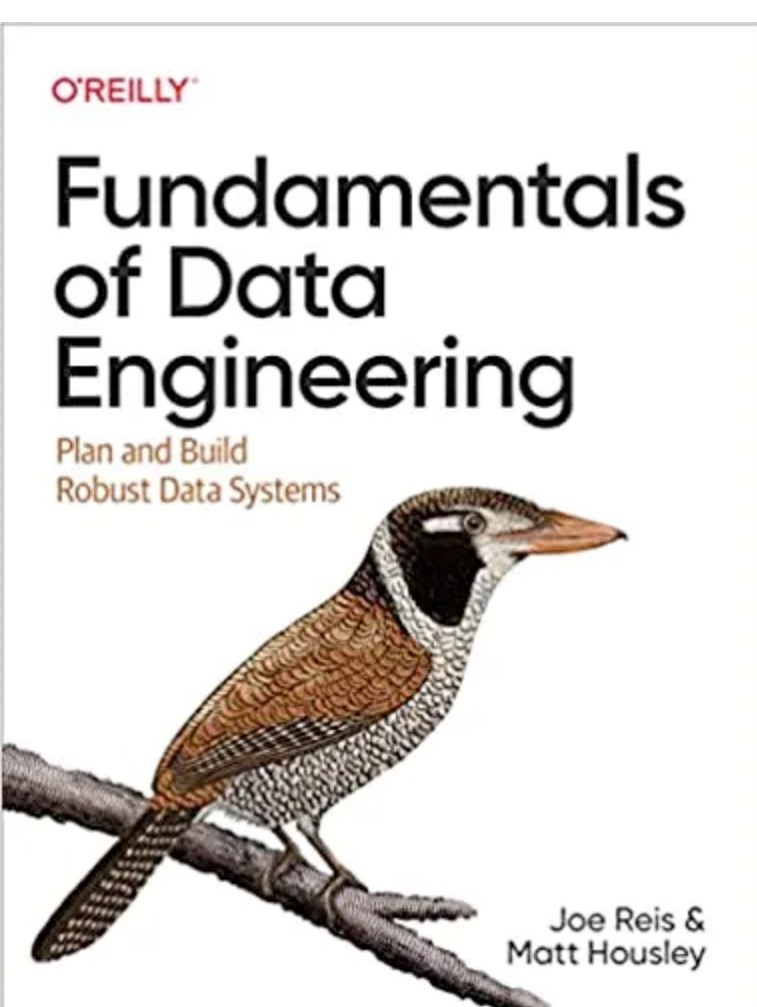
Khairina · Follow  
Published in Data Engineer Things · 14 min read · May 14

796

In this post, I will share my key learning points from the “*Fundamentals of Data Engineering*” book by Joe Reis and Matt Housley. O'Reilly published the book in June 2022. As an aspiring data engineer, I continually seek resources to broaden my understanding and improve my skills. I purchased the book on [Amazon](#) during the Black Friday sale for ~USD 35 and finished reading the book on 26 April 2023.

My notes on the book have blown up to 56 pages. Although my primary goal isn't to understand everything or retain every detail, I enjoy the exposure to new concepts that I can delve deeper into when needed. What I love most about reading technical books is exploring new concepts in-depth and gaining a **more profound understanding** of the field. It also inspired me to reflect on my work experience in data and try to connect the dots between theory and practice.

## What to expect from Fundamentals of Data Engineering book: A brief overview



“*Fundamentals of Data Engineering*” is a comprehensive guide to data engineering concepts and practices. The book is not focused on a particular tool, technology, or platform used in data engineering. It focuses on the **fundamental concepts** behind data engineering. It allows you to understand the “why” behind the “how.” This well-written book offers a comprehensive overview of data engineering, making it an excellent choice for beginners and experienced professionals.

Here are my key learning points:

### Chapter 1 — Data Engineering Described

This chapter provides a basic introduction to what data engineering is and the data engineering lifecycle (Generation — Ingestion — Transformation — Serving), which are covered more extensively in the other chapters.

#### Data engineers build a solid foundation

Data engineering provides data to downstream stakeholders such as data scientists. Data scientists spend most of their time on the bottom 3 parts: gathering, cleaning, and processing data. They may not be trained to write production-grade systems.

It is crucial for data engineers to build a solid foundation in these 3 parts before tackling...

Create an account to read the full story.

The author made this story available to Medium members only.  
If you're new to Medium, create a new account to read this story on us.

- Sign up with Google
- Sign up with Facebook
- Sign up with email

Already have an account? [Sign in](#)

796

Written by Khairina

413 Followers · Writer for Data Engineer Things

Data @ Coca-Cola | Singapore | Data Engineering | Analytics

Follow

Post

More from Khairina and Data Engineer Things

Project 1: Spotify ETL using Python and AWS

In this post, I will share the steps taken and key learnings from the pipeline project in th...

11 min read · May 17

231 2

10 Things I Learned from Reading Fundamentals of Data Engineering

After two enriching years as a Data Engineer, I finally had the chance to dive into...

10 min read · Aug 2

1.4K 15

Column lineage is out: AI is in

Why column-level lineage will rapidly be replaced by artificial intelligence

7 min read · Dec 19

149 7

Data Pipelines Pocket Reference—Key Learning Points

In this post, I will share my key learning points from the Data Pipelines Pocket Reference...

19 min read · Jul 2

120 2

See all from Khairina

See all from Data Engineer Things

## Recommended from Medium

Vishal Banvaliya

Roadmap for Data Engineers in 2024

Navigating the Data Engineering Landscape in 2024: A Comprehensive Roadmap

4 min read · Dec 22

452 3

Nishchay Agrawal

Data Engineer Walmart Interview Experience

Subscribe My Youtube Channel For Data Engineering Insights & Interview Experience...

9 min read · Sep 21

282 7

Lists

Predictive Modeling w/ Python

20 stories · 734 saves

Practical Guides to Machine Learning

10 stories · 847 saves

General Coding Knowledge

20 stories · 725 saves

Coding & Development

11 stories · 343 saves

Vechtomova Maria in Marvelous MLOps

MLOps roadmap 2024

The MLOps engineer role is different from an ML engineer role. Even though the role varie...

6 min read · Dec 21

706 5

Nnamdi Samuel in Art of Data Engineering

The Meta Database Engineer Capstone Project

My Approach To Tackling The Tasks Involved

9 min read · Oct 3

195

Richard Warepan in ILLUMINATION

How to Break into Data Engineering : 2024

The One & Only Way: Mastering the Fundamentals

11 min read · Dec 11

170 2

Michael Berk in Towards Data Science

1.5 Years of Spark Knowledge in 8 Tips

My learnings from Databricks customer engagements

8 min read · 5 days ago

459 3

See more recommendations

Help Status About Careers Blog Privacy Terms Text to speech Teams

