



॥ त्वं ज्ञानमयो विद्वानमयोऽसि ॥

Who are Data Engineers?

Saptarshi Pyne

Assistant Professor


Department of Computer Science and Engineering
Indian Institute of Technology Jodhpur, Rajasthan, India 342030

CSL4030 Data Engineering Lecture 1

July 31st, 2023

[Jobs on the Rise Reports](#) → United States Report

Jobs on the Rise in 2021

 Insights from LinkedIn

LinkedIn listed 'Data Engineering' as one of the top jobs on the rise

Data science specialists

According to the 2020 U.S. Emerging Jobs Report, data roles, specifically data scientist and data engineer roles, are increasing steadily — reflecting about a 35% average annual growth for both roles.



Top titles

Data Management Analyst, Data Scientist, Data Science Specialist, Data Engineers, and Data Architect



Top skills

Tableau, SQL, R



Top regions

Washington, DC–Baltimore area, Greater Chicago area, Atlanta metropolitan area

Compensations offered to Data Engineers

How much does a Data Engineer make?

Updated Jul 29, 2023

Experience

Industry

All years of Experience

All industries

\$115,398 /yr

Total Pay

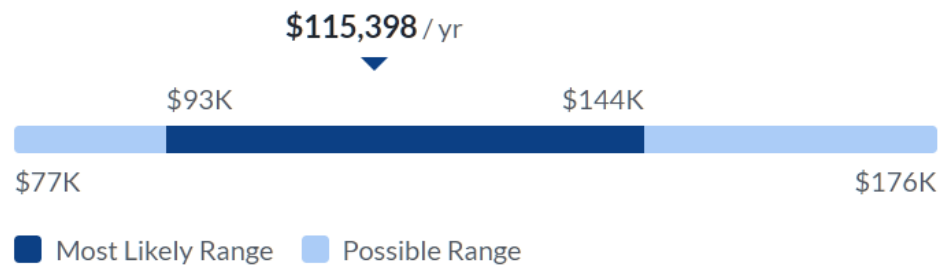
\$98,485 /yr

Base Pay

\$16,913 /yr

Additional Pay

\$121,500 = ₹1 Crore



Total Pay Trajectory For Data Engineer

- \$115,398 /yr
Data Engineer
- \$170,203 /yr
Senior Data Engineer
- \$172,714 /yr
Lead Data Engineer

[See Full Career Path >](#)

Download as data table

So, who are these Data Engineers?

Data Engineers are the people who build systems to manage the flow of data.

So, who are these Data Engineers?

Data Engineers are the people who build systems to manage the flow of data.

How are they different from Data Scientists?

- Data Scientists analyze data to answer a question related to the domain the data belongs to.

Data Engineers vs. Data Scientists: A case study in the domain of Space Research

A Moon rover sending data to the Earth

- Contains millions of sensors for monitoring Moon's soil
- Travels and monitors Moon's soil 24x7
- Sends Terrabytes of data to Earth every second



ISRO Data Engineers

- Where can I store such massive amount of data efficiently?
- How do I make sure that users can perform search and sort operations efficiently on this data?

ISRO Data Scientists

- How does the soil temperature varies from Moon's equator to polar regions?

Who are some famous Data Engineers?

Are there any people who decided to manage the flow of the whole Internet, and make it easily searchable and rankable?

Who are some famous Data Engineers?

Are there any people who decided to manage the flow of the whole Internet, and make it easily searchable and rankable?



Who are some famous Data Engineers? (contd.)

Doug Cutting co-authored Apache Hadoop.
It is an open source software suite for
managing **Big Data**.



Big Data

What is **Big Data**?

Big Data

What is **Big Data**?

- “Big data is **high-volume**, **high-velocity** and **high-variety** information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making.” ~ Gartner

What we discussed today

- Who are Data Engineers?
- How are they different from Data Scientists?
- What is the famous 'Apache Hadoop'?

What we will discuss in the next class

- What is Big Data?
- What are the differences between structured, semi-structured, and unstructured data?
- Can we use traditional data-processing apps with Big Data?

Thank you