



- ✓
- CSV ✓
 - PDF
 - XML
 - JSON, YAML
 - mp4, mp3
 - jpeg, png
 - ORC, RC, AVRO, Parquet
 - txt

- cloud
- Salesforce
- SAP
- Oracle
- Sensor
- Apps. ✓

- Batch
- Real time ✓
- Near Real time.
- Structured
- unstructured
- Semi-structured.

✓ ✓ ✓
E T L → DB, File System

BD ✓

DS ✓

DA/BA ✓

- ① Software Eng → 0 DOE
- ② DOE → 1-2
- ③ Senior Big data Engg → 3-5
- ④ Lead Data Engg → 7-8
- ⑤ Manager/Architect → 10P → 12P

✓ ✓ ✓ ✓
(m2) / ML / CV / NLP

① SDS ✓

② DS

③ SDS

④ LDS

⑤ MDS

① GDA

② DA

③ SOA

④ LDA

⑤ MDA

① Social media

⑦

② Gmail

③ E-commerce

④ Facial Recog → Digit

⑤ Similarity GAN

⑥ Long Trans later