

1. Peran utama seorang data engineer adalah melakukan proses pengambilan data, melakukan data cleaning terhadap data yang error atau kosong, melakukan transformasi data untuk menghasilkan data yang lebih mudah dibaca, menganalisa dan mempublish data menjadi data yang bermakna dan berguna bagi perusahaan maupun organisasi tertentu. Perbedaan antara data engineer dan data analyst adalah data analyst bertugas melakukan analisis terhadap data yang telah melalui berbagai proses yang dilakukan oleh data engineer guna menghasilkan kesimpulan atau informasi yang bermanfaat bagi perusahaan/organisasi.
2. Data analyst dan data engineer sama-sama berkerja dengan data, melakukan pre-processing terhadap data untuk menghasilkan data yang bebas dari noise, melakukan visualisasi data untuk menemukan kesimpulan atau informasi berharga penting dari sebuah data.
3. Proses ETL diawali dengan melakukan extract data dari beberapa source, kemudian data disimpan pada lokasi sementara dan dilakukan transformasi data kemudian data yang telah dimodifikasi akan disimpan didalam data warehouse. Sedangkan untuk proses ETL diawali dengan extract data dari source langsung pada data warehouse, lalu kemudian dilakukan transformasi/perubahan pada data langsung pada data warehouse.