

1. A paragraph on what PaaS, SaaS and IaaS are and the differences between them.

The basic difference between PaaS, SaaS, and IaaS are the different platforms they utilize. IaaS is provided with an API, and you are available to do whatever you want on the infrastructure. IaaS is also a good source of data storage. While in PaaS, “framework, software and tools” has been already prepared by the service provider, which could save related cost for coding (otherwise needs to create by yourself). At last, for SaaS, everything was pre-coded and packed into software, and what we get to use is the software with all the preset functions developed already. Whatever the software could work on, we will be able to use it. So normally speaking, no coding is required. However, if something is not formerly implanted in the system, we cannot do it.

2. A paragraph on the differences between ETL and ELT. Also, list the pros and cons of each in a chart.

E represents Extract, T represents Transformation, and L represents Load. The main difference between them is when to perform the data transformation, before or after loading data. Which is often derived from the final platform we would like to keep our data. If it is a data warehouse in which strict formatting is required, ETL is needed. While in a system like data lake, we could directly put our data into the system without much trouble with formatting issues.

ETL Pros and Cons	ELT Pros and Cons
slow	fast
Well-developed	Under development
Data processed before using	Data ready right after loading (raw though)
Structured data	Unstructured data allowed
Masking available	Masking not available
Good for small datasets	Good for large datasets