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

IaaS, as known as infrastructure as a service, which provides virtualized computing resources over the internet. And for PaaS, it is as known as Platform as a service. It provides developers with framework software and tools needed to build apps and software, such as Apprenda. And lastly, SaaS, as known as software as a service, it make sure that users can directly access and run the software over the internet without downloading like GoogleDrive.

The differences between them are that IaaS users can deploy and run any software, including operating system, databases and application software, on the infrastructure provided by cloud service providers. And the next level of the service is the PaaS, it is similar to IaaS, except that PaaS also provides the operating system and databases, so users can only focus on the applications and data. And finally, SaaS provides the simplest interface for the end users, less control and flexibility than other two services but much easier to operate, and SaaS products do not involve a download or installation, it can saving the end users from needing to manage software updates, all they are responsible for is their own data.

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

ETL stands for Extract, Transform, and Load. ETL is a data integration methodology that will extracts raw data rom sources first, and then transform the data on a secondary processing server, after that, only once the extracted data has been successfully transformed, then we can load it into a target databases. And unlike ETL, after we extract the raw data, we load it directly into a target data warehouse, and then we do all the data transformation.

|  |  |  |
| --- | --- | --- |
|  | ETL | ELT |
| Data size | ETL is better suited for dealing with smaller data sets that require complex transformations. So it is good for OLAP data. | ELT is better suited when dealing with massive amounts of structured and unstructured data, for example NoSql data. |
| Speed | Performed on secondary server. Usually ETL load time is longer than ELT because it is a multistage process. | Data loading happens faster because there is no waiting for transformations and the data only loads one time into the target data system. |
| Costs | Separate servers can create cost issues. | Simplified data stack costs less |
| Data Lake Compatibility | ETL does not have that | ELT have data lake |
| Privacy | Pre-load transformation can eliminate PII (helps for HIPPA). | Direct loading of data requires more privacy safeguards. |
| Maintenance | Secondary processing server adds to the maintenance burden. | With fewer systems, the maintenance burden is reduced. |
|  |  |  |