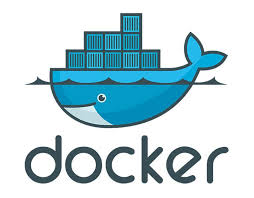
**Introduction of Docker**

# **1.What is Docker ?**



*Docker* is like a **virtual shipping container** for software. Imagine you have an app that you want to run, but it needs specific tools, libraries, or even an operating system to work properly. Instead of worrying about whether your computer has everything the app needs, *Docker* packages everything into a single "*container*" that can run anywhere.

# **2. Data Pipeline**

Another workflow of Data Pipeline :

Environment of Docker for simple data pipeline :

Environment of GCP for simple data pipeline :

# **3. Why we are need the docker for data engineering ?**

1. Consistency Across Environments
2. Efficient Resource Management
3. Reproducibility
4. Compatibility with Modern Cloud Platforms