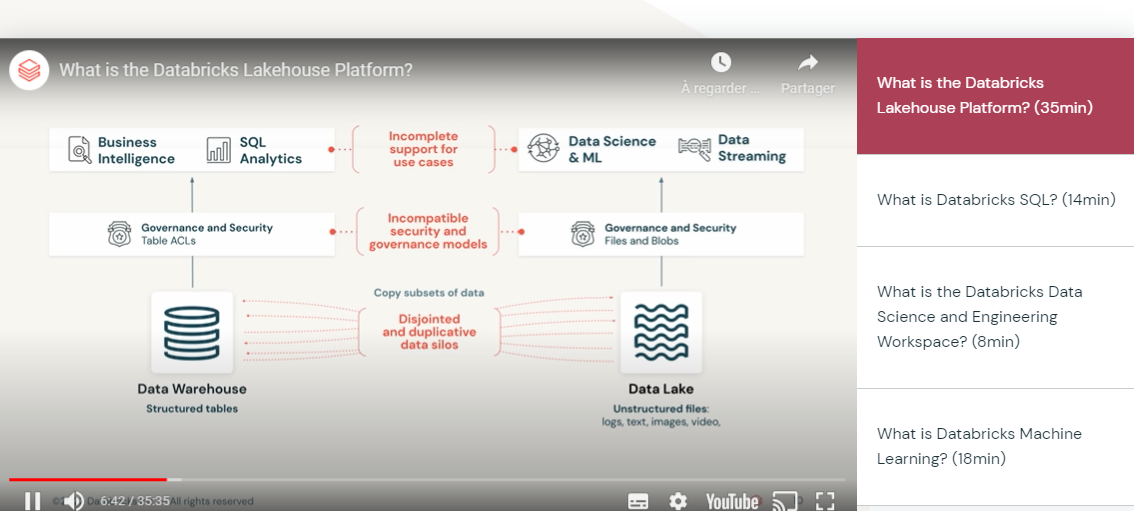
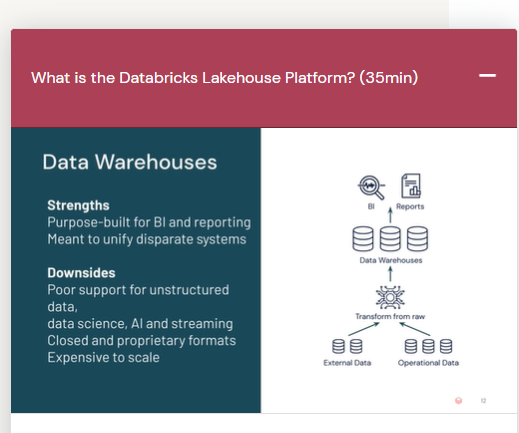
Lakehouse Fundamentals

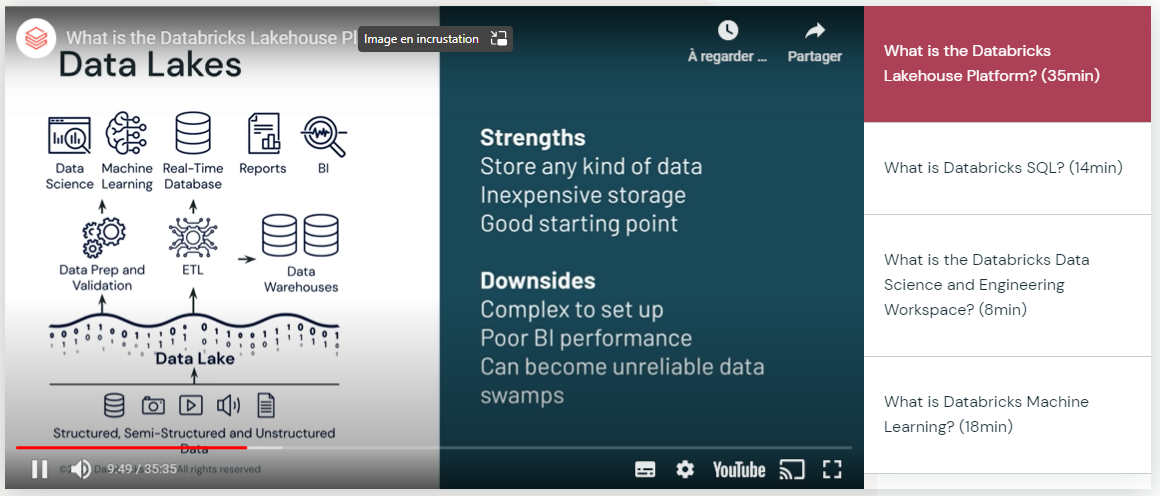
What is the databricks lakehouse platform ?

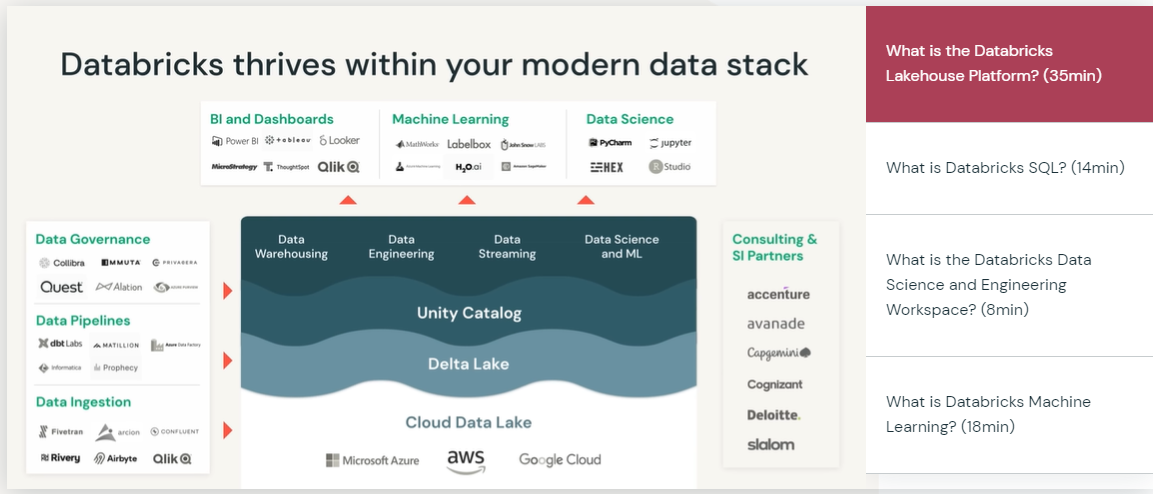
1/ Describe the databricks lakehouse platform

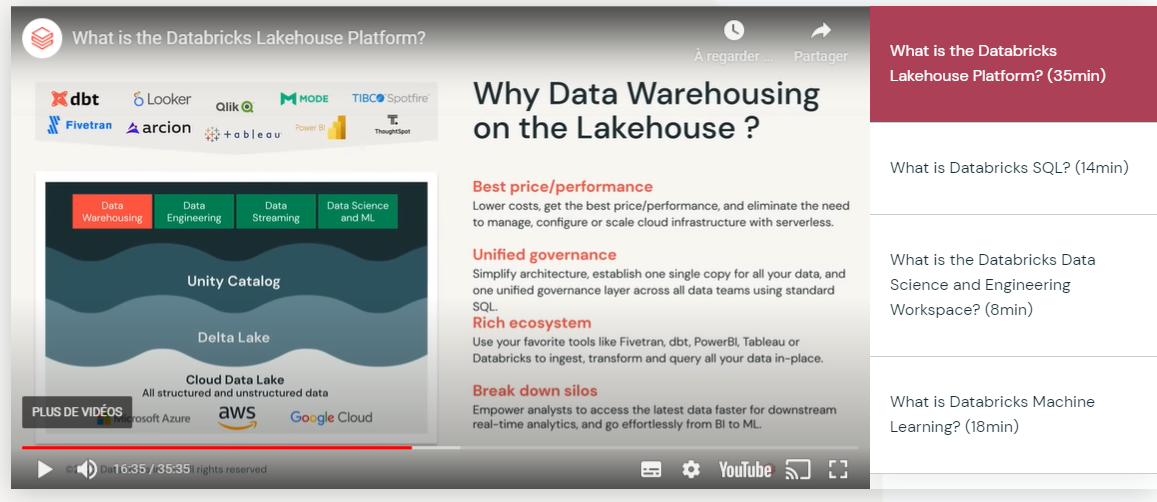
* DataBricks is Data and IA platform
* Pioneer of the data lakehouse
* Gartner recognized leader in both
  + Database Management sustems
  + Data Science and machine learning plafroms
* Creator of highly successful OSS data
  + Delta lake
  + Apache spark
  + MLFlow
* Global adoption
  + Retail
  + Finance ..
* Ingest and explore data + machine learning
* Assosicate datascience and machne
* Web interface :
* Single unified system
* Commit to open source project
* Acces to open source project through platgorm
* 450 partner to gel use platforms
* Allow data science team to collaborate
* Why
  + Big Data and AI challenges
    - 83% of CEO
    - 85 of big
  + Big data high to manage (volume, wide colums
  + Manage type data
  + Kee track of data
  + Many tools , doesn’t work together
  + Avoid multiple stack set up
  + Easy data security and governance (on the same platform
  + 
* Lakehouse paradigm
  + Combined data wahouse and data lake
    - Data warehouse

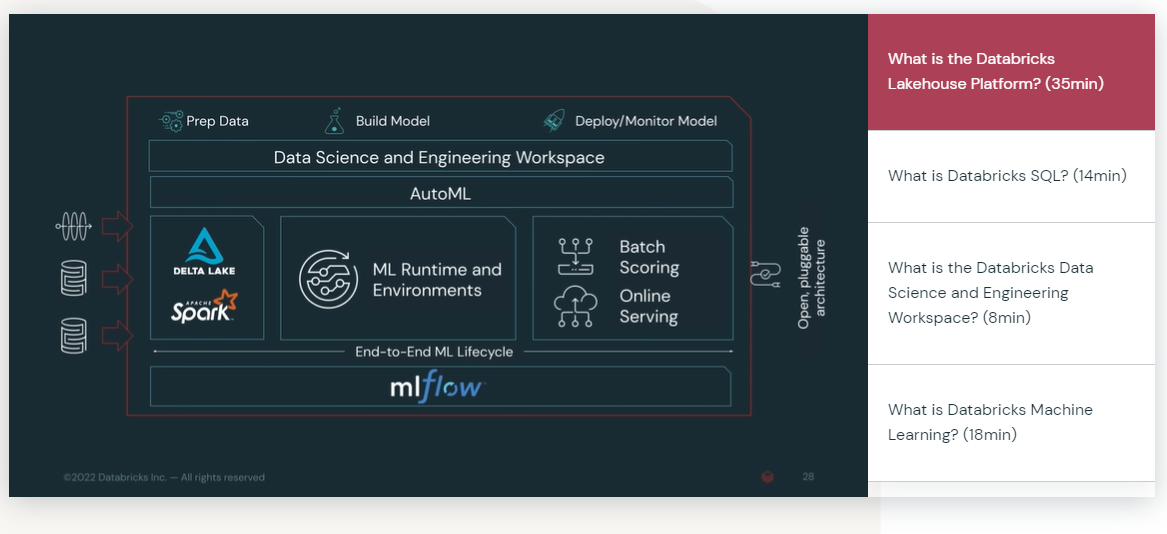


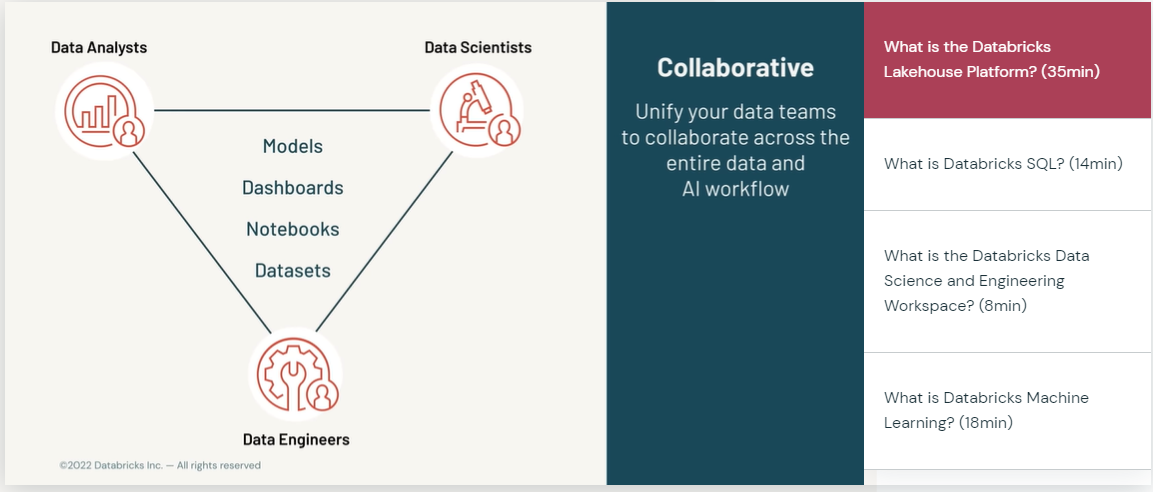
* Data lake (sotre cheap and efficiently)
  + Any nature, any format
  + Decouple data warehouse design, not need to structure and clean the data
  + Take data from datalake => data warehouse > BI
  + Volume can grow quickly
  + Accesing multiple system
  + Most complete and actual data for data science
  + Data schema help to aggregate time to construct
* Lake house
  + Delta lake

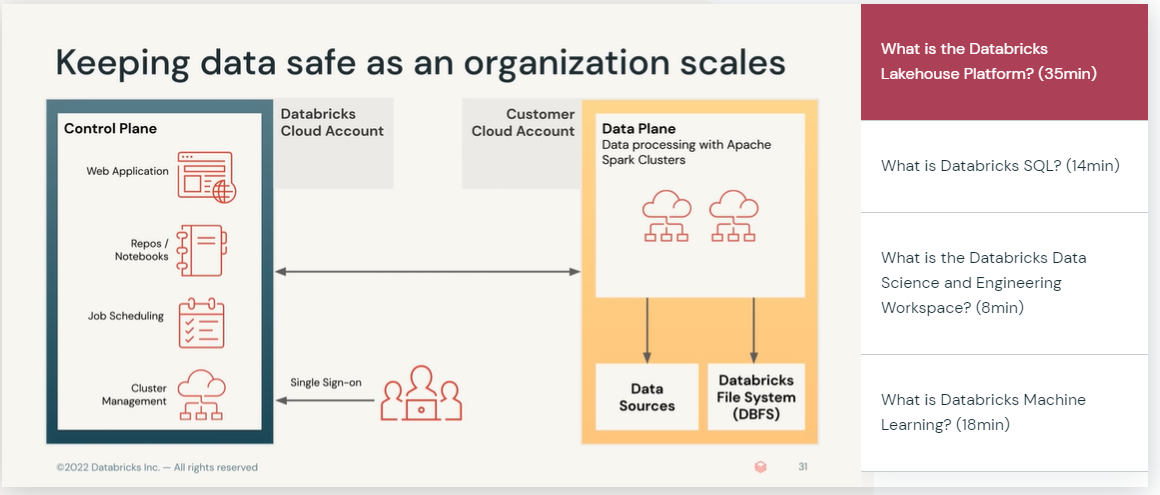


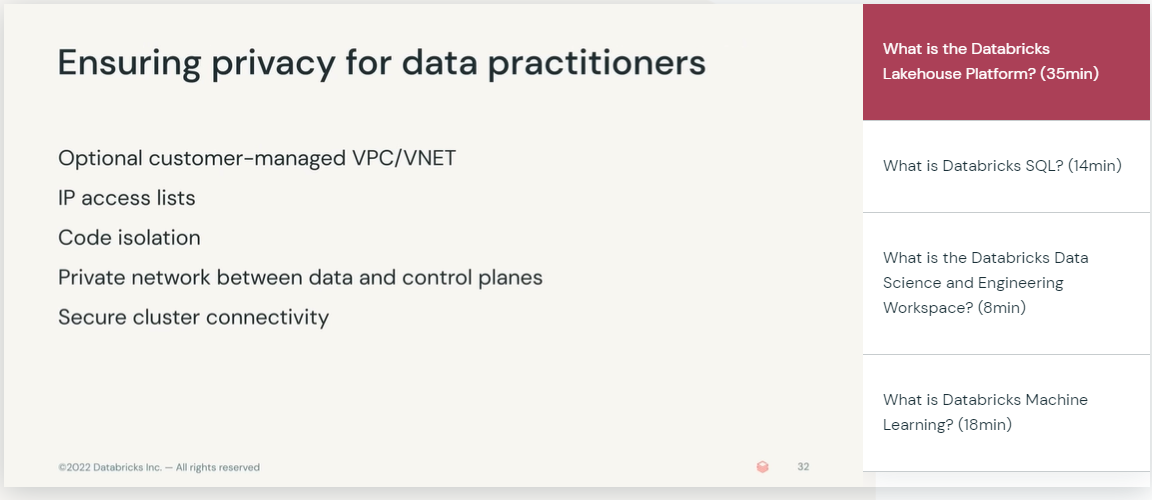


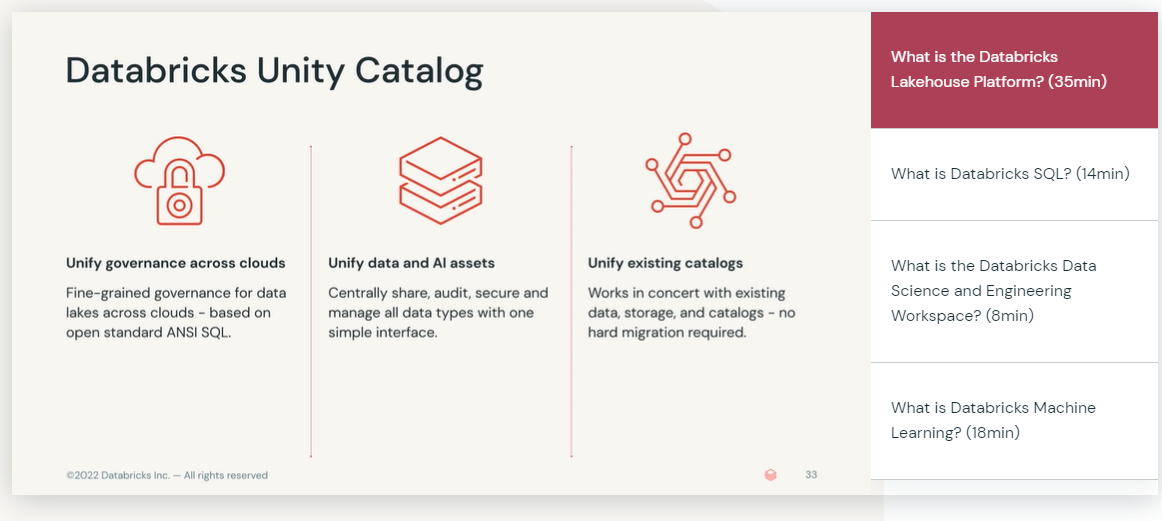


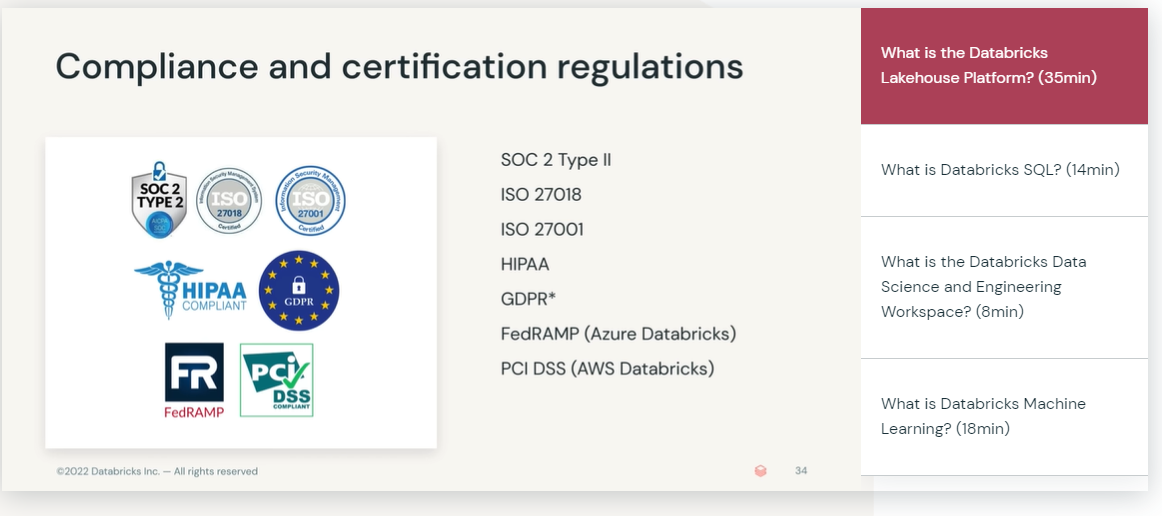














2: explain the origin of the lakehouse data management paradimg

3. outiline fundamental challenfes related to managing and using data

4. describe security features of the databricks lakehouse platform

5 Gie example of orgainsation that benefited from using the databricks lakehouse platform

