Hadoop:
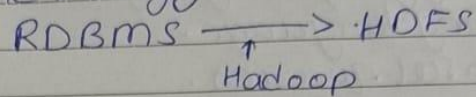
It is an open source framework used for storage and processing of large datasets which is free to use and modify.

$$RDBMS \xrightarrow[\text{Hadoop}]{} HDFS$$

Tehnologies & Tools. ⇒ Hive, Spark, Pig, MapReduce.
Hadoop forms interface linking between RDBMS & HDFS.
Symbol → Yellow . elephant
Founded by → GFS (Google File System).

Greating a database.
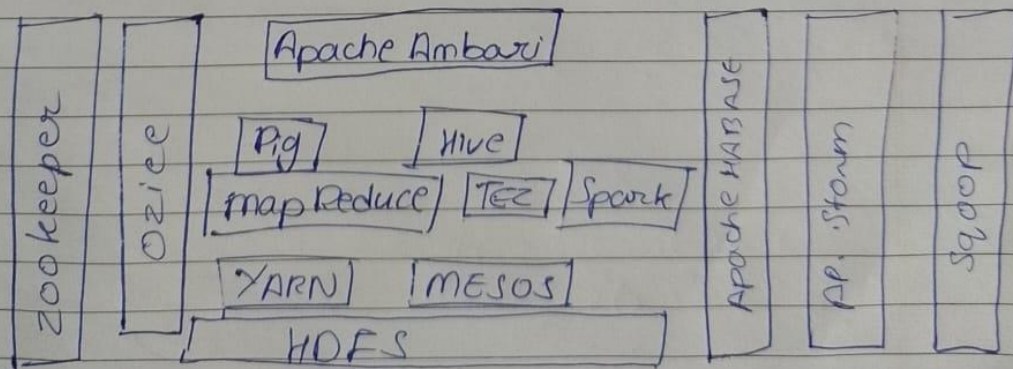1) CREATE DATABASE Employee;
↳ CREATE TABLE Employee;
Table Schema → Rows / Columns.
Primary Key → Unique identity.

Hadoop . Architecture.

Zookeeper	Oziee	Apache Ambari			Apache HABASE	Ap. Storm	Sqoop
		Pig	Hive				
		map Reduce	TEZ	Spark			
		YARN	MESOS				
		HDFS					

↳ YERN = Yet Another Resource Negotiation (to manage resources).

↳ MAP = Distribution of dataset.

↳ Reducer = Aggregating together all the data.

↳ Pig = Scripting language.

↳ Hive = Run sql queries.

↳ TIZ = Graphical representation.

↳ SPARK = For scaler programing language.

↳ APACHE = Exporting data analytics.

↳ APACHE STORM = Real-time processing.

↳ Sqoop = sql. ——> Hdfs.

↳ Ozice = Helps in scheduling.

↳ Zookeeper = keeps the record about the states present the system.

↳ Apache Ambari = Gives interface.