**關聯式資料庫簡介及優缺點：**

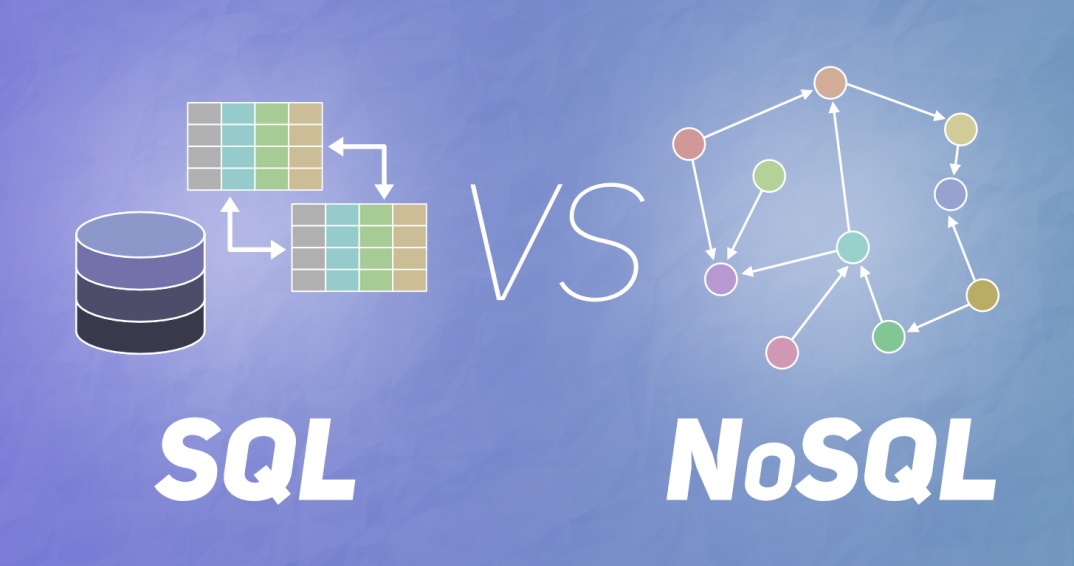
關聯式資料庫為兩個或兩個以上的資料表組合而成，而資料表中的資料是透過行與列的二元形式表現，資料表經過正規化後根據其共同的欄位值進行連結並儲存結構化的資料。

**優點：**

1. 節省重複輸入的時間與儲存空間。
2. 確保異動資料的一致性及完整性。
3. 結構化的資料更易於使用及管理。
4. 有統一的SQL語法。
5. ACID標準保證了資料的穩定、安全和可預測性。

**缺點：**

1. 效能容易隨著資料庫的擴大(表格增加)而快速下降。

****

圖片來源：[https](https://www.geeksforgeeks.org/sql-vs-nosql-which-one-is-better-to-use/) : [//www.geeksforgeeks.org/sql-vs-nosql-which-one-is-better-to-use/](https://www.geeksforgeeks.org/sql-vs-nosql-which-one-is-better-to-use/)

**非關聯式資料庫簡介及優缺點：**

因應個人電腦、網路的普及，使用者利用社群網站或應用程式快速產出大量非結構化資料，則善於處理高速及大量產生資料且不需要固定結構的NoSQL資料庫隨之興起。

**優點：**

1. 擁有鍵值、文件和圖形等資料類型，針對不同資料的存取更加高效。
2. 因每份資料皆為獨立，所以容易分散資料及擴充儲存空間。
3. 善於處理高速及大量產生的資料且不需要固定結構。

**缺點：**

1. 各家供應商常擁有自己的語法，缺乏標準化。
2. 因放寬ACID標準，所以並不保證資料的精確度，也可能導致資料遺失。

# 非關聯式資料庫介紹（Apache Flink）：

****

Apache Flink是由Apache軟體基金會於2011年5月開發，Apache Flink是一個框架分散式系統，用於在無邊界和有邊界資料流上進行有狀態的計算。

**架構：**

任何資料都可以成為一種資料流，而資料流分為無界資料流和有界資料流，無界流會無止境的產生資料，這些資料需要立即且持續處理，有界流可以在資料蒐集結束後再進行處理、計算，Apache Flink擅長處理此兩種資料流，精確的時間控制和狀態化能夠運行任何無界流的應用，有界流則由一些專為固定大小資料集所設計的算法和資料結構進行內部處理，並產生優異效能。

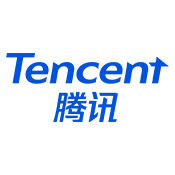
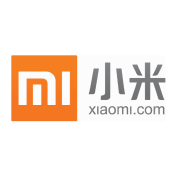
Apache Flink是分散式系統，能夠部署在各種集群環境，例如[Hadoop YARN](https://hadoop.apache.org/docs/stable/hadoop-yarn/hadoop-yarn-site/YARN.html)、[Apache Mesos](https://mesos.apache.org/)和[Kubernetes](https://kubernetes.io/)，同時也支援單機運行，且能夠應用於任意規模。

**應用：**

Apache Flink有三應用場景，分別為Data Pipelie、Data Analytics、Data Driven；時間是資料流處理應用中重要的組成部分，時間有Event-time和Processing-time兩種模式，而Watermark Support和Late Data Handling為特殊事件及其衍生事件。

Apache Flink的API分為三層，由上至下分為SQL / Table API、DataStream API、ProcessFunction，每種API針對不同場景在簡潔性及表達力有不同的側重。

**知名使用企業、網站：**

** **